

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number
WO 00/73801 A2

(51) International Patent Classification¹: G01N 33/574, C12Q 1/68, A61K 39/00, 39/395, 35/14, 48/00, C12N 15/63, 5/10, C07K 14/47, 7/04, 16/18 (72) Inventor; and
(75) Inventor/Applicant (for US only): OBATA, Yuichi [JP/JP]; Chikusa-ku, Nagoya 464 (JP).

(21) International Application Number: PCT/US00/14749

(74) Agent: VAN AMSTERDAM, John, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

(22) International Filing Date: 26 May 2000 (26.05.2000)

(81) Designated States (national): AU, CA, CN, JP, KR, US.

(25) Filing Language: English

(84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

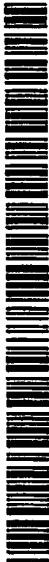
(30) Priority Data:
60/136,526 28 May 1999 (28.05.1999) US
60/153,454 10 September 1999 (10.09.1999) US

Published:

— Without international search report and to be republished upon receipt of that report.

(71) Applicant (for all designated States except US): LUDWIG INSTITUTE FOR CANCER RESEARCH [CH/US]; 605 Third Avenue, New York, NY 10158 (US).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



A2

(54) Title: BREAST, GASTRIC AND PROSTATE CANCER ASSOCIATED ANTIGENS AND USES THEREFOR

(57) Abstract: Cancer associated antigens have been identified by autologous antibody screening of libraries of nucleic acids expressed in breast, gastric and prostate cancer cells using antisera from cancer patients. The invention relates to nucleic acids and encoded polypeptides which are cancer associated antigens expressed in patients afflicted with cancer. The invention provides, *inter alia*, isolated nucleic acid molecules, expression vectors containing those molecules and host cells transfected with those molecules. The invention also provides isolated proteins and peptides, antibodies to those proteins and peptides and cytotoxic T lymphocytes which recognize the proteins and peptides. Fragments of the foregoing including functional fragments and variants also are provided. Kits containing the foregoing molecules additionally are provided. The molecules provided by the invention can be used in the diagnosis, monitoring, research, or treatment of conditions characterized by the expression of one or more cancer associated antigens.

WO 00/73801 A2

**BREAST, GASTRIC AND PROSTATE CANCER ASSOCIATED
ANTIGENS AND USES THEREFOR**

Field of the Invention

5 The invention relates to nucleic acids and encoded polypeptides which are cancer associated antigens expressed in patients afflicted with breast, gastric or prostate cancer. The invention also relates to agents which bind the nucleic acids or polypeptides. The nucleic acid molecules, polypeptides coded for by such molecules and peptides derived therefrom, as well as related antibodies and cytolytic T lymphocytes, are useful, *inter alia*, in diagnostic and
10 therapeutic contexts.

Background of the Invention

15 The mechanism by which T cells recognize foreign materials has been implicated in cancer. A number of cytolytic T lymphocyte (CTL) clones directed against autologous melanoma antigens, testicular antigens, and melanocyte differentiation antigens have been described. In many instances, the antigens recognized by these clones have been characterized.

20 The use of autologous CTLs for identifying tumor antigens requires that the target cells which express the antigens can be cultured *in vitro* and that stable lines of autologous CTL clones which recognize the antigen-expressing cells can be isolated and propagated. While this approach has worked well for melanoma antigens, other tumor types, such as epithelial cancers including breast and colon cancer, have proved refractory to the approach.

25 More recently another approach to the problem has been described by Sahin et al. (*Proc. Natl. Acad. Sci. USA* 92:11810-11813, 1995). According to this approach, autologous antisera are used to identify immunogenic protein antigens expressed in cancer cells by screening expression libraries constructed from tumor cell cDNA. Antigen-encoding clones so identified have been found to elicit a high-titer humoral immune response in the patients from which the antisera were obtained. Such a high-titer IgG response implies helper T cell recognition of the detected antigen. These tumor antigens can then be screened for the presence of MHC/HLA class I and class II motifs and reactivity with CTLs.

30 Since the individual tumor antigens presently known may be expressed only in a fraction of tumors, the availability of additional tumor antigens would significantly enlarge the

- 2 -

proportion of patients who are potentially eligible for therapeutic interventions. Thus there presently is a need for additional tumor antigens for development of therapeutics and diagnostics applicable to a greater number of cancer patients having various cancers.

The invention is elaborated upon further in the disclosure which follows.

5

Summary of the Invention

Autologous antibody screening has now been applied to breast, gastric and prostate cancer using antisera from cancer patients. Numerous cancer associated antigens have been identified. The invention provides, *inter alia*, isolated nucleic acid molecules, expression vectors containing those molecules and host cells transfected with those molecules. The invention also provides isolated proteins and peptides, antibodies to those proteins and peptides and CTLs which recognize the proteins and peptides. Fragments including functional fragments and variants of the foregoing also are provided. Kits containing the foregoing molecules additionally are provided. The foregoing can be used in the diagnosis, monitoring, research, or treatment of conditions characterized by the expression of one or more cancer associated antigens.

Prior to the present invention, only a handful of cancer associated genes had been identified in the past 20 years. The invention involves the surprising discovery of several genes, some previously known and some previously unknown, which are expressed in individuals who have cancer. These individuals all have serum antibodies against the proteins (or fragments thereof) encoded by these genes. Thus, abnormally expressed genes are recognized by the host's immune system and therefore can form a basis for diagnosis, monitoring and therapy.

The invention involves the use of a single material, a plurality of different materials and even large panels and combinations of materials. For example, a single gene, a single protein encoded by a gene, a single functional fragment thereof, a single antibody thereto, etc. can be used in methods and products of the invention. Likewise, pairs, groups and even panels of these materials and optionally other cancer associated antigen genes and/or gene products can be used for diagnosis, monitoring and therapy. The pairs, groups or panels can involve 2, 3, 4, 5 or more genes, gene products, fragments thereof or agents that recognize such materials. A plurality of such materials are not only useful in monitoring, typing, characterizing and diagnosing cells abnormally expressing such genes, but a plurality of such

- 3 -

materials can be used therapeutically. An example of the use of a plurality of such materials for the prevention, delay of onset, amelioration, etc. of cancer cells, which express or will express such genes prophylactically or acutely. Any and all combinations of the genes, gene products, and materials which recognize the genes and gene products can be tested and identified for use according to the invention. It would be far too lengthy to recite all such combinations; those skilled in the art, particularly in view of the teaching contained herein, will readily be able to determine which combinations are most appropriate for which circumstances.

As will be clear from the following discussion, the invention has *in vivo* and *in vitro* uses, including for therapeutic, diagnostic, monitoring and research purposes. One aspect of the invention is the ability to fingerprint a cell expressing a number of the genes identified according to the invention by, for example, quantifying the expression of such gene products. Such fingerprints will be characteristic, for example, of the stage of the cancer, the type of the cancer, or even the effect in animal models of a therapy on a cancer. Cells also can be screened to determine whether such cells abnormally express the genes identified according to the invention.

The invention, in one aspect, is a method of diagnosing a disorder characterized by expression of a cancer associated antigen precursor coded for by a nucleic acid molecule. The method involves the steps of contacting a biological sample isolated from a subject with an agent that specifically binds to the nucleic acid molecule, an expression product thereof, or a fragment of an expression product thereof complexed with an MHC, preferably an HLA, molecule, wherein the nucleic acid molecule is a NA Group 1 nucleic acid molecule, and determining the interaction between the agent and the nucleic acid molecule, the expression product or fragment of the expression product as a determination of the disorder.

In one embodiment the agent is selected from the group consisting of (a) a nucleic acid molecule comprising NA Group 1 nucleic acid molecules or a fragment thereof, (b) a nucleic acid molecule comprising NA Group 3 nucleic acid molecules or a fragment thereof, (c) a nucleic acid molecule comprising NA Group 5 nucleic acid molecules or a fragment thereof, (d) an antibody that binds to an expression product, or a fragment thereof, of NA group 1 nucleic acids, (e) an antibody that binds to an expression product, or a fragment thereof, of NA group 3 nucleic acids, (f) an antibody that binds to an expression product, or a fragment thereof, of NA group 5 nucleic acids, (g) and agent that binds to a complex of an

- 4 -

MHC, preferably HLA, molecule and a fragment of an expression product of a NA Group 1 nucleic acid, (h) an agent that binds to a complex of an MHC, preferably HLA, molecule and a fragment of an expression product of a NA group 3 nucleic acid, and (i) an agent that binds to a complex of an MHC, preferably HLA, molecule and a fragment of an expression product of a NA Group 5 nucleic acid.

The disorder may be characterized by expression of a plurality of cancer associated antigen precursors. Thus the methods of diagnosis may include use of a plurality of agents, each of which is specific for a different human cancer associated antigen precursor (including at least one of the cancer associated antigen precursors disclosed herein), and wherein said plurality of agents is at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9 or at least 10 such agents. Any of the diagnostic methods disclosed herein can be applied sequentially over time to permit determination of the prognosis or progression (or regression) of the disorder.

In each of the above embodiments the agent may be specific for a human cancer associated antigen precursor, including the breast, gastric and prostate cancer associated antigen precursors disclosed herein.

In another aspect the invention is a method for determining regression, progression or onset of a condition characterized by expression of abnormal levels of a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule. The method involves the steps of monitoring a sample, from a subject who has or is suspected of having the condition, for a parameter selected from the group consisting of (i) the protein, (ii) a peptide derived from the protein, (iii) an antibody which selectively binds the protein or peptide, and (iv) cytolytic T cells specific for a complex of the peptide derived from the protein and an MHC molecule, as a determination of regression, progression or onset of said condition. In one embodiment the sample is a body fluid, a body effusion or a tissue.

In another embodiment the step of monitoring comprises contacting the sample with a detectable agent selected from the group consisting of (a) an antibody which selectively binds the protein of (i), or the peptide of (ii), (b) a protein or peptide which binds the antibody of (iii), and (c) a cell which presents the complex of the peptide and MHC molecule of (iv). In a preferred embodiment the antibody, the protein, the peptide or the cell is labeled with a radioactive label or an enzyme. The sample in a preferred embodiment is assayed for the peptide. Preferably samples are isolated from tissue or bodily fluids of the subject at

- 5 -

sequential time points, and the samples are assayed as a determination of the regression, progression or onset of the condition from a first sequential time point to a second sequential time point.

- According to another embodiment the nucleic acid molecule is one of the following:
- 5 NA Group 3 molecule or a NA Group 5 molecule. In yet another embodiment the protein is a plurality of proteins, the parameter is a plurality of parameters, each of the plurality of parameters being specific for a different one of the plurality of proteins.

- The invention in another aspect is a pharmaceutical preparation for a human subject. The pharmaceutical preparation includes an agent which when administered to the subject
- 10 enriches selectively the presence of complexes of an HLA molecule and a human cancer associated antigen, and a pharmaceutically acceptable carrier, wherein the human cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule which comprises a NA Group 1 molecule. In one embodiment the nucleic acid molecule is a NA Group 3 nucleic acid molecule.
 - 15 The agent in one embodiment comprises a plurality of agents, each of which enriches selectively in the subject complexes of an HLA molecule and a different human cancer associated antigen. Preferably the plurality is at least two, at least three, at least four or at least 5 different such agents.

- In another embodiment the agent is selected from the group consisting of (1) an isolated polypeptide comprising the human cancer associated antigen, or a functional variant thereof, (2) an isolated nucleic acid operably linked to a promoter for expressing the isolated polypeptide, or functional variant thereof, (3) a host cell expressing the isolated polypeptide, or functional variant thereof, and (4) isolated complexes of the polypeptide, or functional variants thereof, and an HLA molecule.

- 25 The agent may be a cell expressing an isolated polypeptide. In one embodiment the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof. In another embodiment the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof, and wherein the cell expresses an HLA molecule that binds the polypeptide.
- 30 The cell can express one or both of the polypeptide and HLA molecule recombinantly. In preferred embodiments the cell is nonproliferative. In yet another embodiment the agent is at least two, at least three, at least four or at least five different polypeptides, each representing a

different human cancer associated antigen or functional variant thereof.

The agent in one embodiment is a PP Group 2 polypeptide. In other embodiments the agent is a PP Group 3 polypeptide or a PP Group 4 polypeptide.

In an embodiment each of the pharmaceutical preparations described herein also includes an adjuvant.

According to another aspect the invention, a composition is provided which includes an isolated agent that binds selectively a PP Group 1 polypeptide. In separate embodiments the agent binds selectively to a polypeptide selected from the following: a PP Group 2 polypeptide, a PP Group 3 polypeptide, a PP Group 4 polypeptide, and a PP Group 5 polypeptide. In other embodiments, the agent is a plurality of different agents that bind selectively at least two, at least three, at least four, or at least five different such polypeptides. In each of the above described embodiments the agent may be an antibody.

In another aspect the invention is a composition of matter composed of a conjugate of the agent of the above-described compositions of the invention and a therapeutic or diagnostic agent. Preferably the conjugate is of the agent and a therapeutic or diagnostic that is an antineoplastic.

The invention in another aspect is a pharmaceutical composition which includes an isolated nucleic acid molecule selected from the group consisting of: (1) NA Group 1 molecules, and (2) NA Group 2 molecules, and a pharmaceutically acceptable carrier. In one embodiment the isolated nucleic acid molecule comprises a NA Group 3 or NA Group 4 molecule. In another embodiment the isolated nucleic acid molecule comprises at least two isolated nucleic acid molecules coding for two different polypeptides, each polypeptide comprising a different cancer associated antigen.

Preferably the pharmaceutical composition also includes an expression vector with a promoter operably linked to the isolated nucleic acid molecule. In another embodiment the pharmaceutical composition also includes a host cell recombinantly expressing the isolated nucleic acid molecule.

According to another aspect of the invention a pharmaceutical composition is provided. The pharmaceutical composition includes an isolated polypeptide comprising a PP Group 1 or a PP Group 2 polypeptide, and a pharmaceutically acceptable carrier. In one embodiment the isolated polypeptide comprises a PP Group 3 or a PP Group 4 polypeptide.

In another embodiment the isolated polypeptide comprises at least two different

polypeptides, each comprising a different cancer associated antigen at least one of which is encoded by a NA group 1 molecule as disclosed herein. In separate embodiments the isolated polypeptides are selected from the following: breast cancer polypeptides or HLA binding fragments thereof and gastric cancer polypeptides or HLA binding fragments thereof.

5 In an embodiment each of the pharmaceutical compositions described herein also includes an adjuvant.

Another aspect the invention is an isolated nucleic acid molecule comprising a NA Group 3 molecule. Another aspect the invention is an isolated nucleic acid molecule comprising a NA Group 4 molecule.

10 The invention in another aspect is an isolated nucleic acid molecule selected from the group consisting of (a) a fragment of a nucleic acid selected from the group of nucleic acid molecules consisting of SEQ ID Nos:1-593, of sufficient length to represent a sequence unique within the human genome, and identifying a nucleic acid encoding a human cancer associated antigen precursor, (b) complements of (a), provided that the fragment includes a
15 sequence of contiguous nucleotides which is not identical to any sequence selected from the sequence group consisting of (1) sequences having the GenBank accession numbers of Table 1 and other sequences publicly available as of the filing date of this application, (2) complements of (1), and (3) fragments of (1) and (2). Preferably the unique fragments are fragments of a nucleic acid selected from the group of nucleic acid molecules consisting of
20 SEQ ID NOS:12, 15, 34-59, 61, 62, 83-95, 186, 190-205, 297, 327-332, and 335-352.

In one embodiment the sequence of contiguous nucleotides is selected from the group consisting of: (1) at least two contiguous nucleotides nonidentical to the sequences in Table 1, (2) at least three contiguous nucleotides nonidentical to the sequences in Table 1, (3) at least four contiguous nucleotides nonidentical to the sequences in Table 1, (4) at least five
25 contiguous nucleotides nonidentical to the sequences in Table 1, (5) at least six contiguous nucleotides nonidentical to the sequences in Table 1, or (6) at least seven contiguous nucleotides nonidentical to the sequences in Table 1.

In another embodiment the fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18
30 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

- 8 -

In yet another embodiment the molecule encodes a polypeptide which, or a fragment of which, binds a human HLA receptor (e.g., class I or class II) or a human antibody.

Another aspect of the invention is an expression vector comprising an isolated nucleic acid molecule of the invention described above operably linked to a promoter.

5 According to one aspect the invention is an expression vector comprising a nucleic acid operably linked to a promoter, wherein the nucleic acid is a NA Group 1 or Group 2 molecule. In another aspect the invention is an expression vector comprising a NA Group 1 or Group 2 molecule and a nucleic acid encoding an MHC, preferably HLA, molecule.

10 In yet another aspect the invention is a host cell transformed or transfected with an expression vector of the invention described above.

15 In another aspect the invention is a host cell transformed or transfected with an expression vector comprising an isolated nucleic acid molecule of the invention described above operably linked to a promoter, or an expression vector comprising a nucleic acid operably linked to a promoter, wherein the nucleic acid is a NA Group 1 or 2 molecule and further comprising a nucleic acid encoding HLA.

20 According to another aspect of the invention an isolated polypeptide encoded by the isolated nucleic acid molecules of the invention, described above, is provided. These include PP Group 1-5 polypeptides. The invention also includes a fragment of the polypeptide which is immunogenic. In one embodiment the fragment, or a portion of the fragment, binds HLA or a human antibody.

25 The invention includes in another aspect an isolated fragment of a human cancer associated antigen precursor which, or a portion of which, binds HLA or a human antibody, wherein the precursor is encoded by a nucleic acid molecule that is a NA Group 1 molecule. In one embodiment the fragment is part of a complex with HLA. In another embodiment the fragment is between 8 and 12 amino acids in length. In another embodiment the invention includes an isolated polypeptide comprising a fragment of the polypeptide of sufficient length to represent a sequence unique within the human genome and identifying a polypeptide that is a human cancer associated antigen precursor.

30 According to another aspect of the invention a kit for detecting the presence of the expression of a cancer associated antigen precursor is provided. The kit includes a pair of isolated nucleic acid molecules each of which consists essentially of a molecule selected from the group consisting of (a) a 12-32 nucleotide contiguous segment of the nucleotide sequence

of any of the NA Group 1 molecules and (b) complements of (a), wherein the contiguous segments are nonoverlapping. In one embodiment the pair of isolated nucleic acid molecules is constructed and arranged to selectively amplify an isolated nucleic acid molecule that is a NA Group 3 molecule. Preferably, the pair amplifies a human NA Group 3 molecule.

5 According to another aspect of the invention a method for treating a subject with a disorder characterized by expression of a human cancer associated antigen precursor is provided. The method includes the step of administering to the subject an amount of an agent, which enriches selectively in the subject the presence of complexes of an HLA molecule and a human cancer associated antigen, effective to ameliorate the disorder, wherein the human
10 cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule selected from the group consisting of (a) a nucleic acid molecule comprising NA group 1 nucleic acid molecules, (b) a nucleic acid molecule comprising NA group 3 nucleic acid molecules, (c) a nucleic acid molecule comprising NA group 5 nucleic acid molecules.

15 In one embodiment the disorder is characterized by expression of a plurality of human cancer associated antigen precursors and wherein the agent is a plurality of agents, each of which enriches selectively in the subject the presence of complexes of an HLA molecule and a different human cancer associated antigen. Preferably the plurality is at least 2, at least 3, at least 4, or at least 5 such agents.

20 In another embodiment the agent is an isolated polypeptide selected from the group consisting of PP Group 1, PP Group 2, PP Group 3, PP Group 4, and PP group 5 polypeptides.

In yet another embodiment the disorder is cancer.

According to another aspect the invention is a method for treating a subject having a condition characterized by expression of a cancer associated antigen precursor in cells of the
25 subject. The method includes the steps of (i) removing an immunoreactive cell containing sample from the subject, (ii) contacting the immunoreactive cell containing sample to the host cell under conditions favoring production of cytolytic T cells against a human cancer associated antigen which is a fragment of the precursor, (iii) introducing the cytolytic T cells to the subject in an amount effective to lyse cells which express the human cancer associated
30 antigen, wherein the host cell is transformed or transfected with an expression vector comprising an isolated nucleic acid molecule operably linked to a promoter, the isolated nucleic acid molecule being selected from the group of nucleic acid molecules consisting of

- 10 -

NA Group 1, NA Group 2, NA Group 3, NA Group 4, NA Group 5.

In one embodiment the host cell recombinantly expresses an HLA molecule which binds the human cancer associated antigen. In another embodiment the host cell endogenously expresses an HLA molecule which binds the human cancer associated antigen.

5 The invention includes in another aspect a method for treating a subject having a condition characterized by expression of a cancer associated antigen precursor in cells of the subject. The method includes the steps of (i) identifying a nucleic acid molecule expressed by the cells associated with said condition, wherein said nucleic acid molecule is a NA Group 1 molecule (ii) transfecting a host cell with a nucleic acid molecule selected from the group 10 consisting of (a) the nucleic acid molecule identified, (b) a fragment of the nucleic acid molecule identified which includes a segment coding for a cancer associated antigen, (c) deletions, substitutions or additions to (a) or (b), and (d) degenerates of (a), (b), or (c); (iii) culturing said transfected host cells to express the transfected nucleic acid molecule, and; (iv) introducing an amount of said host cells or an extract thereof to the subject effective to 15 increase an immune response against the cells of the subject associated with the condition.

Preferably, the antigen is a human antigen and the subject is a human.

In one embodiment the method also includes the step of (a) identifying an MHC molecule which presents a portion of an expression product of the nucleic acid molecule, wherein the host cell expresses the same MHC molecule as identified in (a) and wherein the 20 host cell presents an MHC binding portion of the expression product of the nucleic acid molecule.

In another embodiment the method also includes the step of treating the host cells to render them non-proliferative.

25 In yet another embodiment the immune response comprises a B-cell response or a T cell response. Preferably the response is a T-cell response which comprises generation of cytolytic T-cells specific for the host cells presenting the portion of the expression product of the nucleic acid molecule or cells of the subject expressing the human cancer associated antigen.

In another embodiment the nucleic acid molecule is a NA Group 3 molecule.

30 Another aspect of the invention is a method for treating or diagnosing or monitoring a subject having a condition characterized by expression of an abnormal amount of a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule. The method includes the

step of administering to the subject an antibody which specifically binds to the protein or a peptide derived therefrom, the antibody being coupled to a therapeutically useful agent, in an amount effective to treat the condition.

In one embodiment the antibody is a monoclonal antibody. Preferably the monoclonal antibody is a chimeric antibody or a humanized antibody.

In another aspect the invention is a method for treating a condition characterized by expression in a subject of abnormal amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule. The method involves the step of administering to a subject at least one of the pharmaceutical compositions of the invention described above in an amount effective to prevent, delay the onset of, or inhibit the condition in the subject. In one embodiment the condition is cancer. In another embodiment the method includes the step of first identifying that the subject expresses in a tissue abnormal amounts of the protein.

The invention in another aspect is a method for treating a subject having a condition characterized by expression of abnormal amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule. The method includes the steps of (i) identifying cells from the subject which express abnormal amounts of the protein; (ii) isolating a sample of the cells; (iii) cultivating the cells, and (iv) introducing the cells to the subject in an amount effective to provoke an immune response against the cells.

In one embodiment the method includes the step of rendering the cells non-proliferative, prior to introducing them to the subject.

In another aspect the invention is a method for treating a pathological cell condition characterized by abnormal expression of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule. The method includes the step of administering to a subject in need thereof an effective amount of an agent which inhibits the expression or activity of the protein.

In one embodiment the agent is an inhibiting antibody which selectively binds to the protein and wherein the antibody is a monoclonal antibody, a chimeric antibody, a humanized antibody or a fragment thereof. In another embodiment the agent is an antisense nucleic acid molecule which selectively binds to the nucleic acid molecule which encodes the protein. In yet another important embodiment the nucleic acid molecule is a NA Group 3 nucleic acid molecule.

The invention includes in another aspect a composition of matter useful in stimulating

- 12 -

an immune response to a plurality of proteins encoded by nucleic acid molecules that are NA Group 1 molecules. The composition is a plurality of peptides derived from the amino acid sequences of the proteins, wherein the peptides bind to one or more MHC molecules presented on the surface of the cells which express an abnormal amount of the protein.

5 In one embodiment at least a portion of the plurality of peptides bind to MHC molecules and elicit a cytolytic response thereto. In another embodiment the composition of matter includes an adjuvant. In another embodiment the adjuvant is a saponin, GM-CSF, or an interleukin. In still another embodiment, the compositions also includes at least one peptide useful in stimulating an immune response to at least one protein which is not encoded
10 by nucleic acid molecules that are NA Group 1 molecules, wherein the at least one peptide binds to one or more MHC molecules.

According to another aspect the invention is an isolated antibody which selectively binds to a complex of: (i) a peptide derived from a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule and (ii) and an MHC molecule to which binds the peptide to
15 form the complex, wherein the isolated antibody does not bind to (i) or (ii) alone.

In one embodiment the antibody is a monoclonal antibody, a chimeric antibody, a humanized antibody or a fragment thereof.

The invention also involves the use of the genes, gene products, fragments thereof, agents which bind thereto, and so on in the preparation of medicaments. A particular
20 medicament is for treating cancers including, e.g., one or more of cancers of the breast, cervix, ovary, prostate, testis, lung, colon, pancreas, stomach, liver, skin (e.g., melanoma), bladder, head and neck, thyroid, blood cells, bone and kidney. Diagnostics for specific cancers and groups of cancers also are envisioned.

In certain preferred embodiment, the nucleic acid molecules are selected from the
25 group consisting of SEQ ID NOs:1-18, and the polypeptides are encoded by these preferred nucleic acid molecules.

Still other embodiments and aspects of the invention will become apparent in connection with the description of the invention which follows.

30

Detailed Description of the Invention

In the above summary and in the ensuing description, lists of sequences are provided. The lists are meant to embrace each single sequence separately, two or more sequences

- 13 -

together where they form a part of the same gene, any combination of two or more sequences which relate to different genes, including and up to the total number on the list, as if each and every combination were separately and specifically enumerated. Likewise, when mentioning fragment size, it is intended that a range embrace the smallest fragment mentioned to the full-length of the sequence (less one nucleotide or amino acid so that it is a fragment), each and every fragment length intended as if specifically enumerated. Thus, if a fragment could be between 10 and 15 in length, it is explicitly meant to mean 10, 11, 12, 13, 14, or 15 in length.

5 The summary and the claims mention antigen precursors and antigens. As used in the summary and in the claims, a precursor is substantially the full-length protein encoded by the coding region of the isolated DNA and the antigen is a peptide which complexes with MHC, preferably HLA, and which participates in the immune response as part of that complex. Such antigens are typically 9 amino acids long, although this may vary slightly.

10 As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments human cancer antigens and human subjects are preferred.

15 The present invention in one aspect involves the cloning of cDNAs encoding human cancer associated antigen precursors using autologous antisera of subjects having breast, gastric or prostate cancer. The sequences of the clones representing genes identified according to the methods described herein are presented in the attached Sequence Listing. Of 20 the foregoing, it can be seen that some of the clones are considered completely novel as no coding regions were found in the databases searched. Other clones are novel but have some nucleotide or amino acid homologies to sequences deposited in databases (mainly EST sequences). Nevertheless, the entire gene sequence was not previously known. In some cases no function was suspected and in other cases, even if a function was suspected, it was not 25 known that the gene was associated with cancer, or with a particular cancer. In all cases, it was not known or suspected that the gene encoded a cancer antigen which reacted with an antibody from autologous sera. Analysis of the clone sequences by comparison to nucleic acid and protein databases determined that still other of the clones surprisingly are closely related to other previously-cloned genes. The sequences of these related genes is also 30 presented in the Sequence Listing. The nature of the foregoing genes as encoding antigens recognized by the immune systems of cancer patients is, of course, unexpected.

The invention thus involves in one aspect cancer associated antigen polypeptides,

genes encoding those polypeptides, functional modifications and variants of the foregoing, useful fragments of the foregoing, as well as diagnostics and therapeutics relating thereto.

Homologs and alleles of the cancer associated antigen nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for cancer associated antigen precursors. Because this application contains so many sequences, the following chart is provided to identify the various groups of sequences discussed in the claims and in the summary:

Nucleic Acid Sequences

- 10 NA Group 1. (a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of a nucleic acid sequence selected from the group consisting of nucleic acid sequences among SEQ ID NOs: 1-593, and which code for a cancer associated antigen precursor,
 - (b) deletions, additions and substitutions which code for a respective cancer associated antigen precursor,
 - (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).
- 20 NA Group 2. Fragments of NA Group 1, which code for a polypeptide which, or a portion of which, binds an MHC molecule to form a complex recognized by an autologous antibody or lymphocyte.
- NA Group 3. The subset of NA Group 1 where the nucleotide sequence is selected from the group consisting of:
 - (a) previously unknown human nucleic acids coding for a human cancer associated antigen precursor, e.g., SEQ ID NOs: 12, 15, 34-59, 61, 62, 83-95, 186, 190-205, 297, 327-332, and 335-352,
 - (b) deletions, additions and substitutions which code for a respective human cancer associated antigen precursor,
 - (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

- 15 -

(d) complements of (a), (b) or (c).

NA Group 4. Fragments of NA Group 3, which code for a polypeptide which, or a portion of which, binds to an MHC molecule to form a complex recognized by an autologous antibody
5 or lymphocyte.

NA Group 5. A subset of NA Group 1, comprising human cancer associated antigens that react with allogeneic cancer antisera.

10 Polypeptide Sequences

- PP Group 1. Polypeptides encoded by NA Group 1.
- PP Group 2. Polypeptides encoded by NA Group 2.
- PP Group 3. Polypeptides encoded by NA Group 3.
- PP Group 4. Polypeptides encoded by NA Group 4.
- 15 PP Group 5. Polypeptides encoded by NA Group 5.

Particularly preferred polypeptides are those recognized by allogeneic sera of cancer patients, but not by non-cancer patient control sera. For example, as shown in the Examples below, polypeptides encoded by SEQ ID NOs:1-18 are recognized only by antibodies in
20 cancer patients antisera.

The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% Bovine Serum Albumin, 2.5 mM NaH₂PO₄(pH7), 0.5% SDS, 2 mM EDTA). SSC is 0.15 M sodium chloride/0.15 M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetraacetic acid. After hybridization, the membrane upon which the DNA is transferred is washed, for example, in 2 x SSC at room temperature and then at 0.1 - 0.5 x SSC/0.1 x SDS at temperatures up to 68°C.
30

- 16 -

There are other conditions, reagents, and so forth which can be used, which result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs
5 and alleles of cancer associated antigen nucleic acids of the invention (e.g., by using lower stringency conditions). The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 80% nucleotide identity
10 and/or at least 90% amino acid identity to the sequences of cancer associated antigen nucleic acid and polypeptides, respectively, in some instances will share at least 90% nucleotide identity and/or at least 95% amino acid identity and in still other instances will share at least 95% nucleotide identity and/or at least 99% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda,
15 Maryland) that can be obtained through the Internet (<ftp://ncbi.nlm.nih.gov/pub/>). Exemplary tools include the BLAST system available at <http://www.ncbi.nlm.nih.gov>, preferably using default settings. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyle-Doolittle hydropathic analysis can be obtained using the MacVector sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic
20 acids also are embraced by the invention.

In screening for cancer associated antigen genes, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film to detect the radioactive signal. In screening for the expression of cancer associated
25 antigen nucleic acids, Northern blot hybridizations using the foregoing conditions can be performed on samples taken from breast, gastric or prostate cancer patients or subjects suspected of having a condition characterized by expression of the cancer associated antigen genes disclosed herein. Amplification protocols such as polymerase chain reaction using primers which hybridize to the sequences presented also can be used for detection of the
30 cancer associated antigen genes or expression thereof.

The breast, gastric and prostate cancer associated genes correspond to SEQ ID Nos:1-593. These sequences represent genes previously known in humans and genes previously

unknown in humans (e.g., SEQ ID NOs:12, 15, 34-59, 61, 62, 83-95, 186, 190-205, 297, 327-332, and 335-352). Preferred breast, gastric and prostate cancer associated antigens for the methods of diagnosis disclosed herein are those which encode polypeptides that react with allogeneic cancer antisera (i.e. NA Group 5). Encoded polypeptides (e.g., proteins), peptides and antisera thereto are also preferred for diagnosis.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulable by standard techniques known to those of ordinary skill in the art. An isolated nucleic acid as used herein is not a naturally occurring chromosome.

As used herein with respect to polypeptides, "isolated" means separated from its native environment and present in sufficient quantity to permit its identification or use. Isolated, when referring to a protein or polypeptide, means, for example: (i) selectively produced by expression cloning or (ii) purified as by chromatography or electrophoresis. Isolated proteins or polypeptides may be, but need not be, substantially pure. The term "substantially pure" means that the proteins or polypeptides are essentially free of other substances with which they may be found in nature or *in vivo* systems to an extent practical and appropriate for their intended use. Substantially pure polypeptides may be produced by techniques well known in the art. Because an isolated protein may be admixed with a pharmaceutically acceptable carrier in a pharmaceutical preparation, the protein may comprise only a small percentage by weight of the preparation. The protein is nonetheless isolated in that it has been separated from the substances with which it may be associated in living systems, i.e. isolated from other proteins.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating cancer associated antigen polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides modified nucleic acid molecules which include additions, substitutions and deletions of one or more nucleotides. In preferred embodiments, these modified nucleic acid molecules and/or the polypeptides they encode retain at least one activity or function of the unmodified nucleic acid molecule and/or the polypeptides, such as antigenicity, enzymatic activity, receptor binding, formation of complexes by binding of peptides by MHC class I and class II molecules, etc. In certain embodiments, the modified nucleic acid molecules encode modified polypeptides, preferably polypeptides having conservative amino acid substitutions as are described elsewhere herein. The modified nucleic acid molecules are structurally related to the unmodified nucleic acid molecules and in preferred embodiments are sufficiently structurally related to the unmodified nucleic acid molecules so that the modified and unmodified nucleic acid molecules hybridize under stringent conditions known to one of skill in the art.

For example, modified nucleic acid molecules which encode polypeptides having single amino acid changes can be prepared. Each of these nucleic acid molecules can have one, two or three nucleotide substitutions exclusive of nucleotide changes corresponding to the degeneracy of the genetic code as described herein. Likewise, modified nucleic acid molecules which encode polypeptides having two amino acid changes can be prepared which have, e.g., 2-6 nucleotide changes. Numerous modified nucleic acid molecules like these will

be readily envisioned by one of skill in the art, including for example, substitutions of nucleotides in codons encoding amino acids 2 and 3, 2 and 4, 2 and 5, 2 and 6, and so on. In the foregoing example, each combination of two amino acids is included in the set of modified nucleic acid molecules, as well as all nucleotide substitutions which code for the 5 amino acid substitutions. Additional nucleic acid molecules that encode polypeptides having additional substitutions (i.e., 3 or more), additions or deletions (e.g., by introduction of a stop codon or a splice site(s)) also can be prepared and are embraced by the invention as readily envisioned by one of ordinary skill in the art. Any of the foregoing nucleic acids or polypeptides can be tested by routine experimentation for retention of structural relation or 10 activity to the nucleic acids and/or polypeptides disclosed herein.

The invention also provides isolated unique fragments of cancer associated antigen nucleic acid sequences or complements thereof. A unique fragment is one that is a 'signature' for the larger nucleic acid. It, for example, is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the cancer associated antigen 15 nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome. Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of the GenBank accession numbers listed in Table 1 or other previously published sequences as of the filing date of the priority documents for sequences 20 listed in a respective priority document or the filing date of this application for sequences listed for the first time in this application which overlap the sequences of the invention.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence 25 other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing 30 PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to

- 20 -

produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the cancer associated antigen polypeptides, useful, for example, in the preparation of antibodies, and in immunoassays. Unique fragments 5 further can be used as antisense molecules to inhibit the expression of cancer associated antigen nucleic acids and polypeptides, particularly for therapeutic purposes as described in greater detail below. Unique fragments also can be used to create chimeric nucleic acid molecule or polypeptide molecules by, for example, joining all or part of the unique fragment to another nucleic acid or polypeptide molecule (homologous or not). For example, the 10 unique fragment may be similar or identical in large part to a known molecule but may have a portion which is nonidentical to the known molecule; the known molecule and the unique fragment can be used to construct a molecule containing in large part the known molecule with the portion unique to the unique fragment added. Other chimeric molecules will be known to one of ordinary skill in the art and can be prepared using standard molecular biology 15 techniques.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of cancer associated antigen sequences and complements thereof will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides (e.g. 12, 13, 20 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 or more bases long), up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last 25 nucleotide (provided the sequence is unique as described above).

Virtually any segment of the polypeptide coding region of novel cancer associated antigen nucleic acids, or complements thereof, that is 25 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the 30 sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

- 21 -

Especially preferred include nucleic acids encoding a series of epitopes, known as "polytopes". The epitopes can be arranged in sequential or overlapping fashion (see, e.g., Thomson et al., *Proc. Natl. Acad. Sci. USA* 92:5845-5849, 1995; Gilbert et al., *Nature Biotechnol.* 15:1280-1284, 1997), with or without the natural flanking sequences, and can be separated by unrelated linker sequences if desired. The polytope is processed to generate individual epitopes which are recognized by the immune system for generation of immune responses.

Thus, for example, peptides derived from a polypeptide having an amino acid sequence encoded by one of the nucleic acid disclosed herein, and which are presented by MHC molecules and recognized by CTL or T helper lymphocytes, can be combined with peptides from one or more other cancer associated antigens (e.g. by preparation of hybrid nucleic acids or polypeptides) to form "polytopes". The two or more peptides (or nucleic acids encoding the peptides) can be selected from those described herein, or they can include one or more peptides of previously known cancer associated antigens. Exemplary cancer associated peptide antigens that can be administered to induce or enhance an immune response are derived from tumor associated genes and encoded proteins including MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A10, MAGE-A11, MAGE-A12, GAGE-1, GAGE-2, GAGE-3, GAGE-4, GAGE-5, GAGE-6, GAGE-7, GAGE-8, GAGE-9, BAGE-1, RAGE-1, LB33/MUM-1, PRAME, NAG, MAGE-B2, MAGE-B3, MAGE-B4, tyrosinase, brain glycogen phosphorylase, Melan-A, MAGE-C1, MAGE-C2, MAGE-C3, MAGE-C4, MAGE-C5, NY-ESO-1, LAGE-1, SSX-1, SSX-2 (HOM-MEL-40), SSX-4, SSX-5, SCP-1 and CT-7. See, for example, PCT application publication no. WO96/10577. Other examples will be known to one of ordinary skill in the art (for example, see Coulie, *Stem Cells* 13:393-403, 1995), and can be used in the invention in a like manner as those disclosed herein. One of ordinary skill in the art can prepare polypeptides comprising one or more peptides and one or more of the foregoing cancer associated peptides, or nucleic acids encoding such polypeptides, according to standard procedures of molecular biology.

Thus polytopes are groups of two or more potentially immunogenic or immune response stimulating peptides which can be joined together in various arrangements (e.g. concatenated, overlapping). The polytope (or nucleic acid encoding the polytope) can be administered in a standard immunization protocol, e.g. to animals, to test the effectiveness of

- 22 -

the polytope in stimulating, enhancing and/or provoking an immune response.

The peptides can be joined together directly or via the use of flanking sequences to form polytopes, and the use of polytopes as vaccines is well known in the art (see, e.g., Thomson et al., *Proc. Acad. Natl. Acad. Sci USA* 92(13):5845-5849, 1995; Gilbert et al., 5 *Nature Biotechnol.* 15(12):1280-1284, 1997; Thomson et al., *J. Immunol.* 157(2):822-826, 1996; Tam et al., *J. Exp. Med.* 171(1):299-306, 1990). For example, Tam showed that polytopes consisting of both MHC class I and class II binding epitopes successfully generated antibody and protective immunity in a mouse model. Tam also demonstrated that polytopes comprising "strings" of epitopes are processed to yield individual epitopes which are 10 presented by MHC molecules and recognized by CTLs. Thus polytopes containing various numbers and combinations of epitopes can be prepared and tested for recognition by CTLs and for efficacy in increasing an immune response.

It is known that tumors express a set of tumor antigens, of which only certain subsets may be expressed in the tumor of any given patient. Polytopes can be prepared which 15 correspond to the different combination of epitopes representing the subset of tumor rejection antigens expressed in a particular patient. Polytopes also can be prepared to reflect a broader spectrum of tumor rejection antigens known to be expressed by a tumor type. Polytopes can be introduced to a patient in need of such treatment as polypeptide structures, or via the use of nucleic acid delivery systems known in the art (see, e.g., Allsopp et al., *Eur. J. Immunol.* 20 26(8):1951-1959, 1996). Adenovirus, pox virus, Ty-virus like particles, adeno-associated virus, plasmids, bacteria, etc. can be used in such delivery. One can test the polytope delivery systems in mouse models to determine efficacy of the delivery system. The systems also can be tested in human clinical trials.

In instances in which a human HLA class I molecule presents tumor rejection antigens 25 derived from cancer associated nucleic acids, the expression vector may also include a nucleic acid sequence coding for the HLA molecule that presents any particular tumor rejection antigen derived from these nucleic acids and polypeptides. Alternatively, the nucleic acid sequence coding for such a HLA molecule can be contained within a separate expression vector. In a situation where the vector contains both coding sequences, the single vector can 30 be used to transfect a cell which does not normally express either one. Where the coding sequences for a cancer associated antigen precursor and the HLA molecule which presents it are contained on separate expression vectors, the expression vectors can be cotransfected.

The cancer associated antigen precursor coding sequence may be used alone, when, e.g. the host cell already expresses a HLA molecule which presents a cancer associated antigen derived from precursor molecules. Of course, there is no limit on the particular host cell which can be used. As the vectors which contain the two coding sequences may be used in any antigen-presenting cells if desired, and the gene for cancer associated antigen precursor can be used in host cells which do not express a HLA molecule which presents a cancer associated antigen. Further, cell-free transcription systems may be used in lieu of cells.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a cancer associated antigen polypeptide, to reduce the expression of cancer associated antigens. This is desirable in virtually any medical condition wherein a reduction of expression of cancer associated antigens is desirable, e.g., in the treatment of cancer. This is also useful for *in vitro* or *in vivo* testing of the effects of a reduction of expression of one or more cancer associated antigens.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon the sequences of nucleic acids encoding breast, gastric or prostate cancer associated antigens, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. For example, a "gene walk" comprising a series of oligonucleotides of 15-30 nucleotides spanning the length of a cancer associated antigen can be prepared, followed by testing for inhibition of cancer associated antigen expression.

Optionally, gaps of 5-10 nucleotides can be left between the oligonucleotides to reduce the number of oligonucleotides synthesized and tested.

In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., *Nature Biotechnol.* 14:840-844, 1996). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In addition, 3'-untranslated regions may be targeted. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although the listed sequences are cDNA sequences, one of ordinary skill in the art may easily derive the genomic DNA corresponding to the cDNA of a cancer associated antigen. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to nucleic acids encoding cancer associated antigens. Similarly, antisense to allelic or homologous cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamidates, carboxymethyl esters and peptides.

10 The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding breast, gastric or prostate cancer associated antigen polypeptides, together with pharmaceutically acceptable carriers.

20 Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art, as further described below.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to, plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate autonomously or integrated in the genome in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β -galactosidase, luciferase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription

of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

5 The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region
10 which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

15 Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a cancer associated antigen polypeptide or fragment or variant thereof.
20 That heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen), which contains an Epstein Barr Virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1 α , which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression

vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant for the expression of an antigen is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

5 Additional vectors for delivery of nucleic acid are provided below.

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of a vector and one or more of the previously discussed cancer associated antigen nucleic acid molecules. Other components may be added, as desired, as long as the previously 10 mentioned nucleic acid molecules, which are required, are included. The invention also includes kits for amplification of a cancer associated antigen nucleic acid, including at least one pair of amplification primers which hybridize to a cancer associated antigen nucleic acid. The primers preferably are 12-32 nucleotides in length and are non-overlapping to prevent formation of "primer-dimers". One of the primers will hybridize to one strand of the cancer 15 associated antigen nucleic acid and the second primer will hybridize to the complementary strand of the cancer associated antigen nucleic acid, in an arrangement which permits amplification of the cancer associated antigen nucleic acid. Selection of appropriate primer pairs is standard in the art. For example, the selection can be made with assistance of a computer program designed for such a purpose, optionally followed by testing the primers for 20 amplification specificity and efficiency.

The invention also permits the construction of cancer associated antigen gene "knock-outs" and transgenic overexpression in cells and in animals, providing materials for studying certain aspects of cancer and immune system responses to cancer.

The invention also provides isolated polypeptides (including whole proteins and 25 partial proteins) encoded by the foregoing cancer associated antigen nucleic acids. Such polypeptides are useful, for example, alone or as fusion proteins to generate antibodies, as components of an immunoassay or diagnostic assay or as therapeutics. Cancer associated antigen polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and 30 eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such

as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

A unique fragment of a cancer associated antigen polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of cancer associated antigens will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 or 12 or more amino acids including each integer up to the full length).

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include interaction with antibodies, interaction with other polypeptides or fragments thereof, selective binding of nucleic acids or proteins, and enzymatic activity. One important activity is the ability to act as a signature for identifying the polypeptide. Another is the ability to complex with HLA and to provoke in a human an immune response. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

The invention embraces variants of the cancer associated antigen polypeptides described above. As used herein, a "variant" of a cancer associated antigen polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a cancer associated antigen polypeptide. Modifications which create a cancer associated antigen variant can be made to a cancer associated antigen polypeptide 1) to reduce or eliminate an activity of a cancer associated antigen polypeptide; 2) to enhance a property of a cancer associated antigen polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) to provide a novel activity or property to a cancer associated antigen polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to an HLA molecule. Modifications to a cancer associated antigen polypeptide are typically made to the nucleic acid which encodes the cancer associated antigen polypeptide, and can include deletions, point

mutations, truncations, amino acid substitutions and additions of amino acids or non-amino acid moieties. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, substitution of L-amino acids with D-amino acids, and the like.

- 5 Modifications also embrace fusion proteins comprising all or part of the cancer associated antigen amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant cancer associated antigen polypeptide according to known methods. One example of such a method is described by Dahiya and Mayo in *Science* 278:82-87, 1997,
- 10 whereby proteins can be designed *de novo*. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiya and Mayo, specific variants of a cancer associated antigen polypeptide can be proposed and tested to determine whether the variant retains a desired conformation. Other computational and computer modeling methods for designing polypeptide mimetics which
- 15 retain activity of the polypeptides described herein, as well as selection methods such as phage display of peptide libraries are known in the art.

In general, variants include cancer associated antigen polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its desired physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a cancer associated antigen polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

- 25 Mutations of a nucleic acid which encode a cancer associated antigen polypeptide preferably preserve the amino acid reading frame of the coding sequence, and preferably do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random 30 mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be

- made to variants (or to non-variant cancer associated antigen polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., *E. coli*, are well known to those of ordinary skill in the art. Still other mutations can be made
- 5 to the noncoding sequences of a cancer associated antigen gene or cDNA clone to enhance expression of the polypeptide. The activity of variants of cancer associated antigen polypeptides can be tested by cloning the gene encoding the variant cancer associated antigen polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the variant cancer associated antigen polypeptide, and testing
- 10 for a functional capability of the cancer associated antigen polypeptides as disclosed herein. For example, the variant cancer associated antigen polypeptide can be tested for reaction with autologous or allogeneic sera as disclosed in the Examples. Preparation of other variant polypeptides may favor testing of other activities, as will be known to one of ordinary skill in the art.
- 15 The skilled artisan will also realize that conservative amino acid substitutions may be made in cancer associated antigen polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e, the variants retain the functional capabilities of the cancer associated antigen polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size
- 20 characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the cancer associated antigen polypeptides include conservative amino acid substitutions of in the amino acid sequences of proteins disclosed herein. Conservative substitutions of amino acids include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T;
- 25 (f) Q, N; and (g) E, D.

For example, upon determining that a peptide derived from a cancer associated antigen polypeptide is presented by an MHC molecule and recognized by CTLs, one can make

- conservative amino acid substitutions to the amino acid sequence of the peptide, particularly at residues which are thought not to be direct contact points with the MHC molecule. For example, methods for identifying functional variants of HLA class II binding peptides are provided in a published PCT application of Strominger and Wucherpfennig (PCT/US96/03182). Peptides bearing one or more amino acid substitutions also can be tested for concordance with known HLA/MHC motifs prior to synthesis using, e.g. the computer program described by D'Amaro and Drijfhout (D'Amaro et al., *Human Immunol.* 43:13-18, 1995; Drijfhout et al., *Human Immunol.* 43:1-12, 1995). The substituted peptides can then be tested for binding to the MHC molecule and recognition by CTLs when bound to MHC.
- These variants can be tested for improved stability and are useful, *inter alia*, in vaccine compositions.

Conservative amino-acid substitutions in the amino acid sequence of cancer associated antigen polypeptides to produce functionally equivalent variants of cancer associated antigen polypeptides typically are made by alteration of a nucleic acid encoding a cancer associated antigen polypeptide. Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a cancer associated antigen polypeptide. Where amino acid substitutions are made to a small unique fragment of a cancer associated antigen polypeptide, such as an antigenic epitope recognized by autologous or allogeneic sera or cytolytic T lymphocytes, the substitutions can be made by directly synthesizing the peptide. The activity of functionally equivalent fragments of cancer associated antigen polypeptides can be tested by cloning the gene encoding the altered cancer associated antigen polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered cancer associated antigen polypeptide, and testing for a functional capability of the cancer associated antigen polypeptides as disclosed herein. Peptides which are chemically synthesized can be tested directly for function, e.g., for binding to antisera recognizing associated antigens.

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits production and/or isolation of the cancer associated antigen protein molecules. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated cancer associated antigen molecules. The

polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause

5 production of the encoded polypeptide. Translation of mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce polypeptide. Those skilled in the art also can readily follow known methods for isolating cancer associated antigen polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

10 The isolation and identification of cancer associated antigen genes also makes it possible for the artisan to diagnose a disorder characterized by expression of cancer associated antigens. These methods involve determining expression of one or more cancer associated antigen nucleic acids, and/or encoded cancer associated antigen polypeptides and/or peptides derived therefrom. In the former situation, such determinations can be carried out via any

15 standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes. In the latter situation, such determinations can be carried out by screening patient antisera for recognition of the polypeptide.

The invention also makes it possible isolate proteins which bind to cancer associated antigens as disclosed herein, including antibodies and cellular binding partners of the cancer

20 associated antigens. Additional uses are described further herein.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from cancer associated antigen polypeptides. A dominant negative polypeptide is an inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or

25 competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative catalytically-inactive kinase which interacts normally with target proteins but does not phosphorylate the target proteins can reduce phosphorylation of

30 the target proteins in response to a cellular signal. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying

promoter binding sites without increasing transcription.

- The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. For example, given the teachings contained herein of breast, gastric and prostate cancer associated antigens, especially those which are similar to known proteins which have known activities, one of ordinary skill in the art can modify the sequence of the cancer associated antigens by site-specific mutagenesis, scanning mutagenesis, partial gene deletion or truncation, and the like.
- See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

The invention also involves agents such as polypeptides which bind to cancer associated antigen polypeptides. Such binding agents can be used, for example, in screening assays to detect the presence or absence of cancer associated antigen polypeptides and complexes of cancer associated antigen polypeptides and their binding partners and in purification protocols to isolated cancer associated antigen polypeptides and complexes of cancer associated antigen polypeptides and their binding partners. Such agents also can be used to inhibit the native activity of the cancer associated antigen polypeptides, for example, by binding to such polypeptides.

The invention, therefore, embraces peptide binding agents which, for example, can be antibodies or fragments of antibodies having the ability to selectively bind to cancer associated antigen polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has

been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')₂ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. See, e.g., U.S. patents 4,816,567, 5,225,539, 5,585,089, 5,693,762 and 5,859,205.

Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')₂, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')₂ fragment antibodies in which

the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which
5 the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves polypeptides of numerous size and type that bind specifically to cancer associated antigen polypeptides, and complexes of both cancer associated antigen polypeptides and their binding partners. These polypeptides may be
10 derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptoids and non-peptide synthetic moieties.
15

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the cancer associated antigen
20 polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the cancer associated antigen polypeptide. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear portion of the sequence that binds to the cancer associated antigen polypeptide can be determined. One can repeat the
25 procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the cancer associated antigen polypeptides. Thus, the cancer associated antigen polypeptides of the invention, or a fragment thereof, can be used to screen peptide libraries, including phage
30 display libraries, to identify and select peptide binding partners of the cancer associated antigen polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for interfering directly with the functioning of

cancer associated antigen and for other purposes that will be apparent to those of ordinary skill in the art.

As detailed herein, the foregoing antibodies and other binding molecules may be used for example to identify tissues expressing protein or to purify protein. Antibodies also may be coupled to specific diagnostic labeling agents for imaging of cells and tissues that express cancer associated antigens or to therapeutically useful agents according to standard coupling procedures. Diagnostic agents include, but are not limited to, barium sulfate, iocetamic acid, iopanoic acid, ipodate calcium, diatrizoate sodium, diatrizoate meglumine, metrizamide, tyropanoate sodium and radiodiagnostics including positron emitters such as fluorine-18 and carbon-11, gamma emitters such as iodine-123, technetium-99m, iodine-131 and indium-111, and nuclides for nuclear magnetic resonance such as fluorine and gadolinium. Other diagnostic agents useful in the invention will be apparent to one of ordinary skill in the art. As used herein, "therapeutically useful agents" include any therapeutic molecule which desirably is targeted selectively to a cell expressing one of the cancer antigens disclosed herein, including antineoplastic agents, radioiodinated compounds, toxins, other cytostatic or cytolytic drugs, and so forth. Antineoplastic therapeutics are well known and include: aminoglutethimide, azathioprine, bleomycin sulfate, busulfan, carmustine, chlorambucil, cisplatin, cyclophosphamide, cyclosporine, cytarabidine, dacarbazine, dactinomycin, daunorubicin, doxorubicin, taxol, etoposide, fluorouracil, interferon- α , lomustine, mercaptopurine, methotrexate, mitotane, procarbazine HCl, thioguanine, vinblastine sulfate and vincristine sulfate. Additional antineoplastic agents include those disclosed in Chapter 52, Antineoplastic Agents (Paul Calabresi and Bruce A. Chabner), and the introduction thereto, 1202-1263, of Goodman and Gilman's "The Pharmacological Basis of Therapeutics", Eighth Edition, 1990, McGraw-Hill, Inc. (Health Professions Division). Toxins can be proteins such as, for example, pokeweed anti-viral protein, cholera toxin, pertussis toxin, ricin, gelonin, abrin, diphtheria exotoxin, or *Pseudomonas* exotoxin. Toxin moieties can also be high energy-emitting radionuclides such as cobalt-60.

In the foregoing methods and compositions, antibodies prepared according to the invention also preferably are specific for the cancer associated antigen/MHC complexes described herein.

When "disorder" is used herein, it refers to any pathological condition where the cancer associated antigens are expressed. An example of such a disorder is cancer, including

breast, gastric and prostate cancer as particular examples.

Samples of tissue and/or cells for use in the various methods described herein can be obtained through standard methods such as tissue biopsy, including punch biopsy and cell scraping, and collection of blood or other bodily fluids by aspiration or other methods.

5 In certain embodiments of the invention, an immunoreactive cell sample is removed from a subject. By "immunoreactive cell" is meant a cell which can mature into an immune cell (such as a B cell, a helper T cell, or a cytolytic T cell) upon appropriate stimulation. Thus immunoreactive cells include CD34⁺ hematopoietic stem cells, immature T cells and immature B cells. When it is desired to produce cytolytic T cells which recognize a cancer 10 associated antigen, the immunoreactive cell is contacted with a cell which expresses a cancer associated antigen under conditions favoring production, differentiation and/or selection of cytolytic T cells; the differentiation of the T cell precursor into a cytolytic T cell upon exposure to antigen is similar to clonal selection of the immune system.

Some therapeutic approaches based upon the disclosure are premised on a response by 15 a subject's immune system, leading to lysis of antigen presenting cells, such as breast, gastric or prostate cancer cells which present one or more cancer associated antigens. One such approach is the administration of autologous CTLs specific to a cancer associated antigen/MHC complex to a subject with abnormal cells of the phenotype at issue. It is within the ability of one of ordinary skill in the art to develop such CTLs *in vitro*. An example of a 20 method for T cell differentiation is presented in International Application number PCT/US96/05607. Generally, a sample of cells taken from a subject, such as blood cells, are contacted with a cell presenting the complex and capable of provoking CTLs to proliferate. The target cell can be a transfectant, such as a COS cell. These transfectants present the desired complex at their surface and, when combined with a CTL of interest, stimulate its 25 proliferation. COS cells are widely available, as are other suitable host cells. Specific production of CTL clones is well known in the art. The clonally expanded autologous CTLs then are administered to the subject.

CTL proliferation can be increased by increasing the level of tryptophan in T cell cultures, by inhibiting enzymes which catabolize tryptophan, such as indoleamine 2,3- 30 dioxygenase (IDO), or by adding tryptophan to the culture. Proliferation of T cells is enhanced by increasing the rate of proliferation and/or extending the number of divisions of the T cells in culture. In addition, increasing tryptophan in T cell cultures also enhances the

lytic activity of the T cells grown in culture.

Another method for selecting antigen-specific CTL clones has recently been described (Altman et al., *Science* 274:94-96, 1996; Dunbar et al., *Curr. Biol.* 8:413-416, 1998), in which fluorogenic tetramers of MHC class I molecule/peptide complexes are used to detect specific 5 CTL clones. Briefly, soluble MHC class I molecules are folded *in vitro* in the presence of β_2 -microglobulin and a peptide antigen which binds the class I molecule. After purification, the MHC/peptide complex is purified and labeled with biotin. Tetramers are formed by mixing the biotinylated peptide-MHC complex with labeled avidin (e.g. phycoerythrin) at a molar ratio of 4:1. Tetramers are then contacted with a source of CTLs such as peripheral blood or 10 lymph node. The tetramers bind CTLs which recognize the peptide antigen/MHC class I complex. Cells bound by the tetramers can be sorted by fluorescence activated cell sorting to isolate the reactive CTLs. The isolated CTLs then can be expanded *in vitro* for use as described herein.

To detail a therapeutic methodology, referred to as adoptive transfer (Greenberg, *J. 15 Immunol.* 136(5): 1917, 1986; Riddel et al., *Science* 257: 238, 1992; Lynch et al, *Eur. J. Immunol.* 21: 1403-1410, 1991; Kast et al., *Cell* 59: 603-614, 1989), cells presenting the desired complex (e.g., dendritic cells) are combined with CTLs leading to proliferation of the CTLs specific thereto. The proliferated CTLs are then administered to a subject with a cellular abnormality which is characterized by certain of the abnormal cells presenting the. 20 particular complex. The CTLs then lyse the abnormal cells, thereby achieving the desired therapeutic goal.

The foregoing therapy assumes that at least some of the subject's abnormal cells present the relevant HLA/cancer associated antigen complex. This can be determined very easily, as the art is very familiar with methods for identifying cells which present a particular 25 HLA molecule, as well as how to identify cells expressing DNA of the pertinent sequences, in this case a cancer associated antigen sequence. Once cells presenting the relevant complex are identified via the foregoing screening methodology, they can be combined with a sample from a patient, where the sample contains CTLs. If the complex presenting cells are lysed by the mixed CTL sample, then it can be assumed that a cancer associated antigen is being presented, 30 and the subject is an appropriate candidate for the therapeutic approaches set forth *supra*.

Adoptive transfer is not the only form of therapy that is available in accordance with the invention. CTLs can also be provoked *in vivo*, using a number of approaches. One

- 40 -

approach is the use of non-proliferative cells expressing the complex. The cells used in this approach may be those that normally express the complex, such as irradiated tumor cells or cells transfected with one or both of the genes necessary for presentation of the complex (i.e. the antigenic peptide and the presenting HLA molecule). Chen et al. (*Proc. Natl. Acad. Sci. USA* 88: 110-114, 1991) exemplifies this approach, showing the use of transfected cells expressing HPV-E7 peptides in a therapeutic regime. Various cell types may be used.

Similarly, vectors carrying one or both of the genes of interest may be used. Viral or bacterial vectors are especially preferred. For example, nucleic acids which encode a cancer associated antigen polypeptide or peptide may be operably linked to promoter and enhancer sequences which direct expression of the cancer associated antigen polypeptide or peptide in certain tissues or cell types. The nucleic acid may be incorporated into an expression vector. Expression vectors may be unmodified extrachromosomal nucleic acids, plasmids or viral genomes constructed or modified to enable insertion of exogenous nucleic acids, such as those encoding cancer associated antigens, as described elsewhere herein. Nucleic acids encoding one or more cancer associated antigens also may be inserted into a retroviral genome, thereby facilitating integration of the nucleic acid into the genome of the target tissue or cell type. In these systems, the gene of interest is carried by a microorganism, e.g., a Vaccinia virus, pox virus, herpes simplex virus, retrovirus or adenovirus, and the materials de facto "infect" host cells. The cells which result present the complex of interest, and are recognized by autologous CTLs, which then proliferate.

A similar effect can be achieved by combining the cancer associated antigen or a stimulatory fragment thereof with an adjuvant to facilitate incorporation into antigen presenting cells *in vivo*. The cancer associated antigen polypeptide is processed to yield the peptide partner of the HLA molecule while a cancer associated antigen peptide may be presented without the need for further processing. Generally, subjects can receive an intradermal injection of an effective amount of the cancer associated antigen. Initial doses can be followed by booster doses, following immunization protocols standard in the art. Preferred cancer associated antigens include those found to react with allogeneic cancer antisera, shown in the examples below.

The invention involves the use of various materials disclosed herein to "immunize" subjects or as "vaccines". As used herein, "immunization" or "vaccination" means increasing or activating an immune response against an antigen. It does not require elimination or

eradication of a condition but rather contemplates the clinically favorable enhancement of an immune response toward an antigen. Generally accepted animal models can be used for testing of immunization against cancer using a cancer associated antigen nucleic acid. For example, human cancer cells can be introduced into a mouse to create a tumor, and one or 5 more cancer associated antigen nucleic acids can be delivered by the methods described herein. The effect on the cancer cells (e.g., reduction of tumor size) can be assessed as a measure of the effectiveness of the cancer associated antigen nucleic acid immunization. Of course, testing of the foregoing animal model using more conventional methods for immunization can include the administration of one or more cancer associated antigen 10 polypeptides or peptides derived therefrom, optionally combined with one or more adjuvants and/or cytokines to boost the immune response. Methods for immunization, including formulation of a vaccine composition and selection of doses, route of administration and the schedule of administration (e.g. primary and one or more booster doses), are well known in the art. The tests also can be performed in humans, where the end point is to test for the 15 presence of enhanced levels of circulating CTLs against cells bearing the antigen, to test for levels of circulating antibodies against the antigen, to test for the presence of cells expressing the antigen and so forth.

As part of the immunization compositions, one or more cancer associated antigens or stimulatory fragments thereof are administered with one or more adjuvants to induce an 20 immune response or to increase an immune response. An adjuvant is a substance incorporated into or administered with antigen which potentiates the immune response. Adjuvants may enhance the immunological response by providing a reservoir of antigen (extracellularly or within macrophages), activating macrophages and stimulating specific sets of lymphocytes. Adjuvants of many kinds are well known in the art. Specific examples of adjuvants include 25 monophosphoryl lipid A (MPL, SmithKline Beecham), a congener obtained after purification and acid hydrolysis of *Salmonella minnesota* Re 595 lipopolysaccharide; saponins including QS21 (SmithKline Beecham), a pure QA-21 saponin purified from *Quillja saponaria* extract; DQS21, described in PCT application WO96/33739 (SmithKline Beecham); QS-7, QS-17, QS-18, and QS-L1 (So et al., *Mol. Cells* 7:178-186, 1997); incomplete Freund's adjuvant; 30 complete Freund's adjuvant; montanide; alum; CpG oligonucleotides (see e.g. Kreig et al., *Nature* 374:546-9, 1995); and various water-in-oil emulsions prepared from biodegradable oils such as squalene and/or tocopherol. Preferably, the peptides are administered mixed with

a combination of DQS21/MPL. The ratio of DQS21 to MPL typically will be about 1:10 to 10:1, preferably about 1:5 to 5:1 and more preferably about 1:1. Typically for human administration, DQS21 and MPL will be present in a vaccine formulation in the range of about 1 μ g to about 100 μ g. Other adjuvants are known in the art and can be used in the invention (see, e.g. Goding, *Monoclonal Antibodies: Principles and Practice*, 2nd Ed., 1986). Methods for the preparation of mixtures or emulsions of peptide and adjuvant are well known to those of skill in the art of vaccination.

Other agents which stimulate the immune response of the subject can also be administered to the subject. For example, other cytokines are also useful in vaccination protocols as a result of their lymphocyte regulatory properties. Many other cytokines useful for such purposes will be known to one of ordinary skill in the art, including interleukin-12 (IL-12) which has been shown to enhance the protective effects of vaccines (see, e.g., *Science* 268: 1432-1434, 1995), GM-CSF and IL-18. Thus cytokines can be administered in conjunction with antigens and adjuvants to increase the immune response to the antigens.

There are a number of immune response potentiating compounds that can be used in vaccination protocols. These include costimulatory molecules provided in either protein or nucleic acid form. Such costimulatory molecules include the B7-1 and B7-2 (CD80 and CD86 respectively) molecules which are expressed on dendritic cells (DC) and interact with the CD28 molecule expressed on the T cell. This interaction provides costimulation (signal 2) to an antigen/MHC/TCR stimulated (signal 1) T cell, increasing T cell proliferation and effector function. B7 also interacts with CTLA4 (CD152) on T cells and studies involving CTLA4 and B7 ligands indicate that the B7-CTLA4 interaction can enhance antitumor immunity and CTL proliferation (Zheng P., et al. *Proc. Natl. Acad. Sci. USA* 95 (11):6284-6289 (1998)).

B7 typically is not expressed on tumor cells so they are not efficient antigen presenting cells (APCs) for T cells. Induction of B7 expression would enable the tumor cells to stimulate more efficiently CTL proliferation and effector function. A combination of B7/IL-6/IL-12 costimulation has been shown to induce IFN-gamma and a Th1 cytokine profile in the T cell population leading to further enhanced T cell activity (Gajewski et al., *J. Immunol.* 154:5637-5648 (1995)). Tumor cell transfection with B7 has been discussed in relation to *in vitro* CTL expansion for adoptive transfer immunotherapy by Wang et al., (*J. Immunol.*, 19:1-8 (1986)). Other delivery mechanisms for the B7 molecule would include nucleic acid (naked DNA) immunization (Kim J., et al. *Nat Biotechnol.*, 15:7:641-646 (1997)) and recombinant viruses

- 43 -

such as adeno and pox (Wendtner et al., *Gene Ther.*, 4:7:726-735 (1997)). These systems are all amenable to the construction and use of expression cassettes for the coexpression of B7 with other molecules of choice such as the antigens or fragment(s) of antigens discussed herein (including polytopes) or cytokines. These delivery systems can be used for induction 5 of the appropriate molecules *in vitro* and for *in vivo* vaccination situations. The use of anti-CD28 antibodies to directly stimulate T cells *in vitro* and *in vivo* could also be considered. Similarly, the inducible co-stimulatory molecule ICOS which induces T cell responses to foreign antigen could be modulated, for example, by use of anti-ICOS antibodies (Hutloff et al., *Nature* 397:263-266, 1999).

10 Lymphocyte function associated antigen-3 (LFA-3) is expressed on APCs and some tumor cells and interacts with CD2 expressed on T cells. This interaction induces T cell IL-2 and IFN-gamma production and can thus complement but not substitute, the B7/CD28 costimulatory interaction (Parra et al., *J. Immunol.*, 158:637-642 (1997), Fenton et al., *J. Immunother.*, 21:2:95-108 (1998)).

15 Lymphocyte function associated antigen-1 (LFA-1) is expressed on leukocytes and interacts with ICAM-1 expressed on APCs and some tumor cells. This interaction induces T cell IL-2 and IFN-gamma production and can thus complement but not substitute, the B7/CD28 costimulatory interaction (Fenton et al., *J. Immunother.*, 21:2:95-108 (1998)). LFA-1 is thus a further example of a costimulatory molecule that could be provided in a 20 vaccination protocol in the various ways discussed above for B7.

Complete CTL activation and effector function requires Th cell help through the interaction between the Th cell CD40L (CD40 ligand) molecule and the CD40 molecule expressed by DCs (Ridge et al., *Nature*, 393:474 (1998), Bennett et al., *Nature*, 393:478 (1998), Schoenberger et al., *Nature*, 393:480 (1998)). This mechanism of this costimulatory 25 signal is likely to involve upregulation of B7 and associated IL-6/IL-12 production by the DC (APC). The CD40-CD40L interaction thus complements the signal 1 (antigen/MHC-TCR) and signal 2 (B7-CD28) interactions.

The use of anti-CD40 antibodies to stimulate DC cells directly, would be expected to enhance a response to tumor antigens which are normally encountered outside of a 30 inflammatory context or are presented by non-professional APCs (tumor cells). In these situations Th help and B7 costimulation signals are not provided. This mechanism might be used in the context of antigen pulsed DC based therapies or in situations where Th epitopes

have not been defined within known cancer antigen precursors.

A cancer associated antigen polypeptide, or a fragment thereof, also can be used to isolate their native binding partners. Isolation of such binding partners may be performed according to well-known methods. For example, isolated cancer associated antigen

5 polypeptides can be attached to a substrate (e.g., chromatographic media, such as polystyrene beads, or a filter), and then a solution suspected of containing the binding partner may be applied to the substrate. If a binding partner which can interact with cancer associated antigen polypeptides is present in the solution, then it will bind to the substrate-bound cancer associated antigen polypeptide. The binding partner then may be isolated.

10 It will also be recognized that the invention embraces the use of the cancer associated antigen cDNA sequences in expression vectors, as well as to transfect host cells and cell lines, be these prokaryotic (e.g., *E. coli*), or eukaryotic (e.g., dendritic cells, B cells, CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells).

Especially useful are mammalian cells such as human, mouse, hamster, pig, goat, primate, etc.

15 They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include keratinocytes, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The expression vectors require that the pertinent sequence, i.e., those nucleic acids described *supra*, be operably linked to a promoter.

The invention also contemplates delivery of nucleic acids, polypeptides or peptides for
20 vaccination. Delivery of polypeptides and peptides can be accomplished according to standard vaccination protocols which are well known in the art. In another embodiment, the delivery of nucleic acid is accomplished by *ex vivo* methods, i.e. by removing a cell from a subject, genetically engineering the cell to include a cancer associated antigen, and reintroducing the engineered cell into the subject. One example of such a procedure is
25 outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction *in vitro* of a functional copy of a gene into a cell(s) of a subject, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous
30 transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. *In vivo* nucleic acid delivery using vectors such as viruses and targeted

- 45 -

liposomes also is contemplated according to the invention.

In preferred embodiments, a virus vector for delivering a nucleic acid encoding a cancer associated antigen is selected from the group consisting of adenoviruses, adeno-associated viruses, poxviruses including vaccinia viruses and attenuated poxviruses, Semliki Forest virus, Venezuelan equine encephalitis virus, retroviruses, Sindbis virus, and Ty virus-like particle. Examples of viruses and virus-like particles which have been used to deliver exogenous nucleic acids include: replication-defective adenoviruses (e.g., Xiang et al., *Virology* 219:220-227, 1996; Eloit et al., *J. Virol.* 7:5375-5381, 1997; Chengalvala et al., *Vaccine* 15:335-339, 1997), a modified retrovirus (Townsend et al., *J. Virol.* 71:3365-3374, 1997), a nonreplicating retrovirus (Irwin et al., *J. Virol.* 68:5036-5044, 1994), a replication defective Semliki Forest virus (Zhao et al., *Proc. Natl. Acad. Sci. USA* 92:3009-3013, 1995), canarypox virus and highly attenuated vaccinia virus derivative (Paoletti, *Proc. Natl. Acad. Sci. USA* 93:11349-11353, 1996), non-replicative vaccinia virus (Moss, *Proc. Natl. Acad. Sci. USA* 93:11341-11348, 1996), replicative vaccinia virus (Moss, *Dev. Biol. Stand.* 82:55-63, 1994), Venezuelan equine encephalitis virus (Davis et al., *J. Virol.* 70:3781-3787, 1996), Sindbis virus (Pugachev et al., *Virology* 212:587-594, 1995), and Ty virus-like particle (Allsopp et al., *Eur. J. Immunol.* 26:1951-1959, 1996). In preferred embodiments, the virus vector is an adenovirus.

Another preferred virus for certain applications is the adeno-associated virus, a double-stranded DNA virus. The adeno-associated virus is capable of infecting a wide range of cell types and species and can be engineered to be replication-deficient. It further has advantages, such as heat and lipid solvent stability, high transduction frequencies in cells of diverse lineages, including hematopoietic cells, and lack of superinfection inhibition thus allowing multiple series of transductions. The adeno-associated virus can integrate into human cellular DNA in a site-specific manner, thereby minimizing the possibility of insertional mutagenesis and variability of inserted gene expression. In addition, wild-type adeno-associated virus infections have been followed in tissue culture for greater than 100 passages in the absence of selective pressure, implying that the adeno-associated virus genomic integration is a relatively stable event. The adeno-associated virus can also function in an extrachromosomal fashion.

In general, other preferred viral vectors are based on non-cytopathic eukaryotic viruses in which non-essential genes have been replaced with the gene of interest. Non-cytopathic

viruses include retroviruses, the life cycle of which involves reverse transcription of genomic viral RNA into DNA with subsequent proviral integration into host cellular DNA.

Adenoviruses and retroviruses have been approved for human gene therapy trials. In general,

the retroviruses are replication-deficient (i.e., capable of directing synthesis of the desired

- 5 proteins, but incapable of manufacturing an infectious particle). Such genetically altered retroviral expression vectors have general utility for the high-efficiency transduction of genes *in vivo*. Standard protocols for producing replication-deficient retroviruses (including the steps of incorporation of exogenous genetic material into a plasmid, transfection of a packaging cell lined with plasmid, production of recombinant retroviruses by the packaging
- 10 cell line, collection of viral particles from tissue culture media, and infection of the target cells with viral particles) are provided in Kriegler, M., "Gene Transfer and Expression, A Laboratory Manual," W.H. Freeman Co., New York (1990) and Murry, E.J. Ed. "Methods in Molecular Biology," vol. 7, Humana Press, Inc., Clifton, New Jersey (1991).

15 Preferably the foregoing nucleic acid delivery vectors: (1) contain exogenous genetic material that can be transcribed and translated in a mammalian cell and that can induce an immune response in a host, and (2) contain on a surface a ligand that selectively binds to a receptor on the surface of a target cell, such as a mammalian cell, and thereby gains entry to the target cell.

20 Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced *in vitro* or *in vivo* in a host. Such techniques include transfection of nucleic acid-CaPO₄ precipitates, transfection of nucleic acids associated with DEAE, transfection or infection with the foregoing viruses including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a 25 vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. Preferred antibodies include antibodies which selectively bind a cancer 30 associated antigen, alone or as a complex with a MHC molecule. Especially preferred are monoclonal antibodies. Where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis

may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

The therapeutics of the invention can be administered by any conventional route, including injection or by gradual infusion over time. The administration may, for example, be oral, intravenous, intraperitoneal, intramuscular, intracavity, subcutaneous, or transdermal.

When cancer associated antigen peptides are used for vaccination, modes of administration which effectively deliver the cancer associated antigen and adjuvant, such that an immune response to the antigen is increased, can be used. For administration of a cancer associated antigen peptide in adjuvant, preferred methods include intradermal, intravenous, intramuscular and subcutaneous administration. Although these are preferred embodiments, the invention is not limited by the particular modes of administration disclosed herein.

Standard references in the art (e.g., *Remington's Pharmaceutical Sciences*, 18th edition, 1990) provide modes of administration and formulations for delivery of immunogens with adjuvant or in a non-adjuvant carrier. When antibodies are used therapeutically, a preferred route of administration is by pulmonary aerosol. Techniques for preparing aerosol delivery systems containing antibodies are well known to those of skill in the art. Generally, such systems should utilize components which will not significantly impair the biological properties of the antibodies, such as the paratope binding capacity (see, for example, Sciarra and Cutie, "Aerosols," in *Remington's Pharmaceutical Sciences*, 18th edition, 1990, pp 1694-1712; incorporated by reference). Those of skill in the art can readily determine the various parameters and conditions for producing antibody aerosols without resort to undue experimentation. When using antisense preparations of the invention, slow intravenous administration is preferred.

The compositions of the invention are administered in effective amounts. An "effective amount" is that amount of a cancer associated antigen composition that alone, or together with further doses, produces the desired response, e.g. increases an immune response to the cancer associated antigen. In the case of treating a particular disease or condition characterized by expression of one or more cancer associated antigens, such as breast, gastric

or prostate cancers, the desired response is inhibiting the progression of the disease. This may involve only slowing the progression of the disease temporarily, although more preferably, it involves halting the progression of the disease permanently. This can be monitored by routine methods or can be monitored according to diagnostic methods of the invention discussed
5 herein. The desired response to treatment of the disease or condition also can be delaying the onset or even preventing the onset of the disease or condition.

Such amounts will depend, of course, on the particular condition being treated, the severity of the condition, the individual patient parameters including age, physical condition, size and weight, the duration of the treatment, the nature of concurrent therapy (if any), the
10 specific route of administration and like factors within the knowledge and expertise of the health practitioner. These factors are well known to those of ordinary skill in the art and can be addressed with no more than routine experimentation. It is generally preferred that a maximum dose of the individual components or combinations thereof be used, that is, the highest safe dose according to sound medical judgment. It will be understood by those of
15 ordinary skill in the art, however, that a patient may insist upon a lower dose or tolerable dose for medical reasons, psychological reasons or for virtually any other reasons.

The pharmaceutical compositions used in the foregoing methods preferably are sterile and contain an effective amount of cancer associated antigen or nucleic acid encoding cancer associated antigen for producing the desired response in a unit of weight or volume suitable
20 for administration to a patient. The response can, for example, be measured by determining the immune response following administration of the cancer associated antigen composition via a reporter system by measuring downstream effects such as gene expression, or by measuring the physiological effects of the cancer associated antigen composition, such as regression of a tumor or decrease of disease symptoms. Other assays will be known to one of
25 ordinary skill in the art and can be employed for measuring the level of the response.

The doses of cancer associated antigen compositions (e.g., polypeptide, peptide, antibody, cell or nucleic acid) administered to a subject can be chosen in accordance with different parameters, in particular in accordance with the mode of administration used and the state of the subject. Other factors include the desired period of treatment. In the event that a
30 response in a subject is insufficient at the initial doses applied, higher doses (or effectively higher doses by a different, more localized delivery route) may be employed to the extent that patient tolerance permits.

In general, for treatments for eliciting or increasing an immune response, doses of cancer associated antigen are formulated and administered in doses between 1 ng and 1 mg, and preferably between 10 ng and 100 µg, according to any standard procedure in the art. Where nucleic acids encoding cancer associated antigen or variants thereof are employed,

5 doses of between 1 ng and 0.1 mg generally will be formulated and administered according to standard procedures. Other protocols for the administration of cancer associated antigen compositions will be known to one of ordinary skill in the art, in which the dose amount, schedule of injections, sites of injections, mode of administration (e.g., intra-tumoral) and the like vary from the foregoing. Administration of cancer associated antigen compositions to

10 mammals other than humans, e.g. for testing purposes or veterinary therapeutic purposes, is carried out under substantially the same conditions as described above.

When administered, the pharmaceutical compositions of the invention are applied in pharmaceutically-acceptable amounts and in pharmaceutically-acceptable preparations. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with

15 the effectiveness of the biological activity of the active ingredients. Such preparations may routinely contain salts, buffering agents, preservatives, compatible carriers, and optionally other therapeutic agents. When used in medicine, the salts should be pharmaceutically acceptable, but non-pharmaceutically acceptable salts may conveniently be used to prepare pharmaceutically-acceptable salts thereof and are not excluded from the scope of the

20 invention. Such pharmacologically and pharmaceutically-acceptable salts include, but are not limited to, those prepared from the following acids: hydrochloric, hydrobromic, sulfuric, nitric, phosphoric, maleic, acetic, salicylic, citric, formic, malonic, succinic, and the like. Also, pharmaceutically-acceptable salts can be prepared as alkaline metal or alkaline earth salts, such as sodium, potassium or calcium salts.

25 A cancer associated antigen composition may be combined, if desired, with a pharmaceutically-acceptable carrier. The term "pharmaceutically-acceptable carrier" as used herein means one or more compatible solid or liquid fillers, diluents or encapsulating substances which are suitable for administration into a human. The term "carrier" denotes an organic or inorganic ingredient, natural or synthetic, with which the active ingredient is

30 combined to facilitate the application. The components of the pharmaceutical compositions also are capable of being co-mingled with the molecules of the present invention, and with each other, in a manner such that there is no interaction which would substantially impair the

- 50 -

desired pharmaceutical efficacy.

The pharmaceutical compositions may contain suitable buffering agents, including: acetic acid in a salt; citric acid in a salt; boric acid in a salt; and phosphoric acid in a salt.

5 The pharmaceutical compositions also may contain, optionally, suitable preservatives, such as: benzalkonium chloride; chlorobutanol; parabens and thimerosal.

The pharmaceutical compositions may conveniently be presented in unit dosage form and may be prepared by any of the methods well-known in the art of pharmacy. All methods include the step of bringing the active agent into association with a carrier which constitutes one or more accessory ingredients. In general, the compositions are prepared by uniformly 10 and intimately bringing the active compound into association with a liquid carrier, a finely divided solid carrier, or both, and then, if necessary, shaping the product.

15 Compositions suitable for oral administration may be presented as discrete units, such as capsules, tablets, lozenges, each containing a predetermined amount of the active compound. Other compositions include suspensions in aqueous liquids or non-aqueous liquids such as a syrup, elixir or an emulsion.

Compositions suitable for parenteral administration conveniently comprise a sterile aqueous or non-aqueous preparation of cancer associated antigen polypeptides or nucleic acids, which is preferably isotonic with the blood of the recipient. This preparation may be formulated according to known methods using suitable dispersing or wetting agents and 20 suspending agents. The sterile injectable preparation also may be a sterile injectable solution or suspension in a non-toxic parenterally-acceptable diluent or solvent, for example, as a solution in 1,3-butane diol. Among the acceptable vehicles and solvents that may be employed are water, Ringer's solution, and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this 25 purpose any bland fixed oil may be employed including synthetic mono- or di-glycerides. In addition, fatty acids such as oleic acid may be used in the preparation of injectables. Carrier formulation suitable for oral, subcutaneous, intravenous, intramuscular, etc. administrations can be found in *Remington's Pharmaceutical Sciences*, Mack Publishing Co., Easton, PA.

30

Examples

Example 1: SEREX screening of breast, gastric and prostate cancer cells

Breast, gastric and prostate cancer cDNA libraries were established, using standard

- 51 -

techniques, and the libraries were screened, using the SEREX methodology described by Sahin et al., *Proc. Natl. Acad. Sci. USA* 92: 11810 (1995), and by Chen et al., *Proc. Natl. Acad. Sci. USA* 94: 1914 (1997), each of which is incorporated by reference in its entirety.

To be specific, total RNA was isolated by homogenizing tumor samples in 4M
5 guanidinium thiocyanate/0.5% sodium N-lauryl sarcosine/25 mM EDTA followed by
centrifugation in 5.7 M CsCl/25 mM sodium acetate/10 µM EDTA at 32,000 rpm. Total
mRNA was removed by passing the sample over an oligo-dT cellulose column. The cDNA
libraries were then constructed by taking 5 µg of mRNA, using standard methodologies to
reverse transcribe the material. Breast cancer libraries were prepared from two different
10 breast cancer patients, referred to as "MT" and "MK". Gastric cancer libraries were prepared
from a gastric cancer patient, referred to as "YS".

The cDNA was used to construct a lambda phage library, and 500 phages were plated
onto XL1-Blue MRF *E. coli*, and incubated for eight hours at 37°C. A nitrocellulose
membrane was then placed on the plate, followed by overnight incubation. The membrane
15 was then washed, four times, with Tris buffered saline (TBS) which contained 0.05% Tween,
and was then immersed in TBS containing 5% non-fat dried milk. After one hour, the
membrane was incubated with conjugates of peroxidase-goat anti human IgG specific for Fc
portions of human antibodies (1:2000, diluted in TBS with 1% BSA). The incubation was
carried out for one hour, at room temperature, and the membrane was then washed three times
20 with TBS. Those clones which produced antibodies were visualized with 0.06%
3,3'diaminobenzidine tetrachloride and 0.015% H₂O₂, in 50 mM Tris (pH 7.5). Any clones
which produced immunoglobulin were marked, and then the membrane was washed, two
further times, with TBS that contained 0.05% Tween, and then twice with "neat" TBS.

The membranes were then incubated in 1:100 diluted patient serum, overnight, at 4°C.
25 The patient serum had been pretreated. Specifically, 5 ml samples were diluted to 10 ml with
TBS containing 1% bovine serum albumin, and 0.02% Na₃N. The serum had been treated to
remove antibodies to bacteriophage, by passing it through a 5 ml Sepharose column, to which
a lysate of *E. coli* Y1090 had been attached, followed by passage over a second column which
had *E. coli* lysate and lysate of *E. coli* infected with lambda bacteriophage. The screening was
30 carried out five times. The samples were then diluted to 50 ml, and kept at -80°C, until used
as described herein.

Following the overnight incubation with the membrane, the membrane was washed

- 52 -

twice with TBS/0.05% Tween 20, and then once with TBS. A further incubation was carried out, using the protocols discussed supra, for the peroxidase labeled antibodies.

The positive clones were then sequenced, using standard techniques. Following comparison of the sequences to information available in data banks, clones were resolved into known and unknown genes. Some clones corresponded to previously identified human proteins and nucleotide sequences, and other clones have not been identified in humans previously, although there were related molecules found in other species. Still other clones represent molecules for which no related sequences were found (most clones contained very short sections (e.g. 25 or fewer nucleotides) that corresponded to portions of unrelated sequences). Some GenBank accession numbers representative of sequences having homology to the cancer associated antigen nucleotide sequences of the invention are presented in Table 1. All of the homologous sequences are accessible in publicly-available databases by reference to the sequences' accession numbers provided in Table 1.

15 Breast cancer clones:

The nucleotide sequences of clones derived from breast cancer patients "MT" and "MK" are presented as SEQ ID NOS:1-205. Polypeptides encoded by open reading frames of the nucleic acid clones are presented as SEQ ID Nos: 594-829; the correspondence between nucleic acid molecules and encoded polypeptides is shown in Table 2.

20

Gastric cancer clones:

The nucleotide sequences of clones derived from gastric cancer patient "YS" are presented as SEQ ID NOS:206-352 (clones beginning with "YS"). Polypeptides encoded by open reading frames of the YS nucleic acid clones are presented as SEQ ID Nos:830-1083; the correspondence between nucleic acid molecules and encoded polypeptides is shown in Table 2.

Prostate cancer clones

The nucleotide sequences of clones derived from prostate cancer patient "ZH" are presented as SEQ ID NOS:353-593(clones beginning with "ZH"). Polypeptides encoded by open reading frames of the ZH nucleic acid clones are presented as SEQ ID Nos:1084-1332; the correspondence between nucleic acid molecules and encoded polypeptides is shown in

Table 2.

Table 1: Sequence homologies (GenBank Accession Numbers)

5	SEQ ID NO. 1 NGO-Br-38 combined NM_006644.1, AF039695.1, AB003334.1, AB003333.1, D86956.1, D67017.1, D67016.1, Z47807.1, NM_013559.1, L40406.1, AB005282.1, AB005281.1, AB005280.1, AB023420.1, X67643.1, NM_008300.1, D85904.1, AC011661.5, AE003611.1, AL109620.4, AC007049.8, AC005992.15, AC007066.4, AC006080.1, AC009155.3, AF222716.1, AC009223.2, AC004251.1, AC002367.1, AL161553.2, AL161539.2, AL117202.1, AL009183.10, Z97336.1, AB006696.1, AI658961.1, AW571648.1, AW474070.1, AA843693.1, AW608075.1, AW470142.1, AW572452.1, AA543054.1, AW385582.1, AI742981.1, AW612980.1, AW612983.1, AI582881.1, AI751853.1, AI378269.1, AI920808.1, AI654608.1, AI819251.1, AI831339.1, AI753470.1, AI312753.1, AI803588.1, AI563996.1, AA232636.1, AW015796.1, AW117974.1, AI668853.1, AA535277.1, AA993280.1, AA632202.1, AA912023.1, AW627645.1, AW027050.1, AI337175.1, AI123280.1, AA761750.1, AW316651.1, AI223412.1, AW771160.1, AA219263.1, AW068948.1, AA482770.1, AA166716.1, AW236067.1, AA166806.1, AA485151.1, AI369932.1, AI250881.1, AA933881.1, AI262020.1, AI751852.1, AI050716.1, H52653.1, AI651186.1, AA678506.1, AA582157.1, AW628153.1, AI493255.1, AW340810.1, AI223825.1, AW837156.1, AA136424.1, AA953645.1, AI582484.1, AI673134.1, AW820299.1, AA394027.1, T58153.1, T36072.1, AW390368.1, F22410.1, AA417317.1, AW200035.1, AA278231.1, AI361237.1, AI288972.1, AA810686.1, AW103624.1, AW604836.1, AA730742.1, AA082043.1, Z20100.1, D58216.1, AI799265.1, D29622.1, AA435594.1, AA233888.1, AA485036.1, AI612928.1, AI630481.1, F07487.1, AA731716.1, AA417255.1, AA804371.1, AA571359.1, AA465183.1, F08794.1, T34783.1, Z41841.1, F03714.1, AL137142.8, AC012569.3, AP001563.1, AC022671.2, AC020999.4, AC011743.3, AP000635.1, AP000610.2, AC008070.3, AC022797.3, AC005506.6, AL096782.3,
25	SEQ ID NO. 2 NGO-Br-39 MK262/T3 5'
30	AF039695.1, AB003334.1, D86956.1, NM_006644.1, AB003333.1, NM_013559.1, D67016.1, L40406.1, Z47807.1, D67017.1, AB005277.1, AB005278.1, NM_011020.1, U23921.1, D49482.1, AB001926.1, NM_014278.1, AB023421.1, L12723.1, AB005279.1, X67643.1, AB005280.1, AF077354.1, NM_008300.1, AB023420.1, D85904.1, AB005281.1, AC024830.1, L08605.1, AC011294.3, AC009424.2, AC022520.2, NM_013393.1, AC019018.7, AF093415.1, AF161311.1, AF136711.1, AE001434.1, AE001433.1, Z49769.1, AC024813.1, AE003645.1, AC011609.9, AC004150.8, AC004801.1, AL163244.2, AP001699.1, AP001605.1, L16771.1, AW820299.1, AW859988.1, AW859943.1, AW604836.1, AW820234.1, AW206874.1, AI094015.1, AA885873.1, AW820232.1, AI702970.1, AW390368.1, AA777564.1, AA580595.1, H91160.1, AA777031.1, AW608075.1, H54657.1, H64019.1, AI658961.1, H63551.1, AA811573.1, AW628153.1, AA749004.1, AI800379.1, W45471.1, AI751852.1, AW385582.1, AI290252.1, AW389335.1, AA953645.1, AW238563.1, AA805016.1, F08794.1, F07487.1, AW631423.1, T63090.1, N84915.1, AW630933.1, AW474070.1, AA166806.1, N84914.1, AI758907.1, AW103624.1, AW571648.1, AA394027.1, AI002886.1, AA094644.1, AW391561.1, AW362751.1, H63595.1, AW609781.1, H54656.1, AW572452.1, W86085.1, AW577563.1, AW820231.1, AW362766.1, AA555929.1, AA555921.1, AA485036.1, AW820224.1, AW391572.1, H91211.1, AW316651.1, AI838486.1, AA571359.1, AJ397361.1, AA334479.1, AW754210.1, AW583074.1, AI760838.1, AW578928.1, AA212025.1, C81194.1, AA645750.1, AW819755.1, AW125594.1, AU080443.1, AA919208.1, AA755774.1, AA615363.1, AA445826.1, AA117945.1, AI337175.1, AW819997.1, AW470142.1, AA626524.1, AA079853.1, W22433.1, T29047.1, AI626242.1, AW839103.1, 45 AU035998.1, AA624532.1, AA572403.1, AA431598.1, AA370218.1, AA571473.1, AW754207.1, AW366794.1, AL137142.8, AC015501.3, AC021286.3, AC006882.2, AC068895.1, AC055115.2, AC013660.4, AL354918.3, AL138763.2, AC010646.4, AC010267.5, AC008642.3, AC008484.3, AC006279.6, AC006278.6, AC016522.4, AC019327.4, AC021435.2, AC011301.4, AF216669.1, AL159973.2, AL034557.7,
50	SEQ ID NO. 3 NGO-Br-39 MK494/T3 5'
55	AF039695.1, AB003334.1, D86956.1, Z47807.1, NM_006644.1, AB003333.1, NM_013559.1, D67016.1, L40406.1, D67017.1, AB005277.1, AB005278.1, AB005276.1, NM_011020.1, U23921.1, D49482.1, AB001926.1, NM_014278.1, AB023421.1, L12723.1, AB005279.1, X67643.1, AF077354.1, NM_008300.1, AB005275.1, AB023420.1, D85904.1, AC009424.2, NM_013393.1, AF093415.1, AC010852.5, AF161311.1, AF136711.1, AC005516.1, AE001434.1, AE001433.1, AC003099.1, Z49769.1, AP001821.1, AC007678.3, AC006403.3, AC024813.1, AE003684.1, AC004668.1, AC004879.1, AC006354.2, AC010183.6, AC005049.2, AC004150.8, AC004801.1, AF049895.1, AF068862.1, AF004739.1, AL162911.1, Z68341.1, AL032629.1, AL023578.1, U41009.1, L16771.1, AI094015.1, AW206874.1, 60 AA777564.1, AA885873.1, AI702970.1, AI800379.1, AA580595.1, AA805016.1, AW631423.1, AA811573.1,

-232-

AL161730.3

SEQ ID NO:549

ZH024/T3

5 AL163203.2, AL050302.2, AL049911.2, AP000026.1, AP000025.1, NM_014915.1, AK001137.1, AB028997.1, AC007617.10, AE003447.1, AJ223186.1, AC015600.6, AE003617.1, AC004739.1, AC006355.3, AC006045.2, U48386.1, AF044083.1, Z70270.1, AL163224.2, AJ250841.1, Z74696.1, U41993.1, D82813.1, AP001679.1, AP001506.1, AP000961.2, AC005522.2, AC007379.2, AC005100.2, AC005479.2, Y18930.1, Z69637.1, AL022395.2, AI951118.1, AI482625.1, AI487533.1, AW000914.1, AI922499.1, AI902224.1, AI871874.1, AA991162.1, AA736439.1, AA661357.1, AA660701.1, C60377.1, AA471702.1, AA219203.1, AA095151.1, AW208236.2, AW519678.1, AV322534.1, AV282871.1, AV382738.1, AI808313.1, AV045752.2, AV043012.2, AI583901.1, AI478844.1, AI311562.1, AI075925.1, AA828186.1, AA701829.1, AA660895.1, AA660377.1, AC067744.2, AL162272.4, AC009401.2, AC027141.1, AC024370.2, AC011254.3, AC012582.3, AC014239.1, AL354836.1, AC062032.2, AC062004.2, AC036209.2, AC022139.3, AC027699.1, AC012248.2, AL353892.2, AP000872.2, AP001133.1, AC022418.3, AC026295.3, 15 AC024918.2, AC021149.4, AC022811.1, AC010988.3, AC016444.2, AC005139.3

SEQ ID NO: 550

ZH024/T7

20 AL033504.3, AC006604.1, AE003524.1, AF016675.2, AC007437.16, AF039039.2, AF047655.1, AC006239.5, AF101310.1, AF098500.1, U80023.1, AF068712.1, AC004219.1, Z81567.1, Z78065.1, AF003144.1, AL163241.2, U50072.1, U40425.1, AP001696.1, AP001597.1, AB015476.1, AC008638.5, AE003639.1, AC002116.1, AC006699.2, AC008102.16, AC005165.1, AC007297.22, AF039904.1, AC007192.1, AF069076.1, AF077542.1, AF098737.1, AF098736.1, AF098735.1, AF050665.1, AF099914.1, AC002375.1, AL161548.2, AF014161.1, AF014160.1, AL021710.1, Z92972.1, Z48543.1, AL121782.9, AL032640.1, AL096800.20, AL163273.2, U80029.1, AL110500.1, U70852.1, 25 AL096772.5, U50065.1, M18370.1, Z80332.1, X06759.1, X06386.1, U15763.1, AP001728.1, AB008822.1, AP001432.1, M55506.1, AP000151.1, D83253.1, AP000009.2, D90195.1, U14163.1, AA601511.1, AA412108.1, AL044891.1, AA393080.1, AI913336.1, AI908240.1, AA489119.1, AI681153.1, AI013642.1, AI145328.1, AA894065.1, AW655882.1, AW654312.1, AW654290.1, AW632141.1, AW481770.1, AW298986.1, AW239842.1, AI594437.1, AI90825.1, AA597263.1, AA409087.1, AL157387.2, AL162272.4, AC022596.4, AC015818.5, AL356259.1, AL162715.3, 30 AC026081.2, AC021384.3, AL139327.13, AC034129.2, AC068536.2, AC040905.2, AC027205.2, AC024004.2, AL356360.1, AL354773.4, AP001996.1, AC044813.3, AC048334.5, AC019230.3, AC011740.2, AC027583.1, AC011009.4, AC021442.1, AC019577.1, AC006903.1, AC006727.1, AC018917.6, AC068778.4, AC010546.3, AC009111.4, AC068376.1, AC064876.1, AC025223.4, AC062034.1, AC026043.3, AC025965.2, AC011714.1, AC009598.2, AC006842.1, AL158089.6, AL137140.5, AL354664.2, AL117348.14, AL031011.20, AL022597.5, 35 AC001486.1, AP000822.1, AL021450.1, Z95393.1

SEQ ID NO: 551

ZH046/T3

40 AL163203.2, AL050302.2, AL049911.2, AP000026.1, AP000025.1, NM_014915.1, AK001137.1, AB028997.1, AC007617.10, AE003447.1, AJ223186.1, AE003494.1, AC004739.1, AC006355.3, AC006045.2, U48386.1, AF044083.1, Z70270.1, AL163224.2, Z74696.1, U41993.1, D82813.1, AP001679.1, AP001506.1, AP000961.2, AC005522.2, AC007379.2, AC006478.2, AC004996.1, AC007099.3, AC005100.2, AC005879.3, AC005331.1, AC005177.1, AC004045.1, Y18930.1, AL133465.30, AL132766.13, AL109985.2, AL022395.2, Z82193.1, AI951118.1, AW000914.1, AI922499.1, AI871874.1, AA991162.1, AU006171.1, AA661357.1, C60377.1, AA471702.1, AA095151.1, AW519678.1, 45 AV282871.1, AI902224.1, AV045752.2, AI311562.1, AI075925.1, AA828186.1, AA701829.1, AC067744.2, AL162272.4, AC007351.16, AC009401.2, AC027141.1, AC024370.2, AC014239.1, AC062032.2, AC036209.2, AC060754.3, AC007131.3, AC061987.1, AC027699.1, AC021187.4, AC013707.2, AC012410.2, AC012542.4, AC012248.2, AC013152.1, AL022284.1, AC062015.2, AC046187.2, AC022418.3, AC008885.3, AC027696.2, AC060232.3, AC004932.2, AC024466.3, AC019259.3, AC016916.4, AC023988.2, AC016444.2, AC005139.3, AL161730.3

50 SEQ ID NO: 552

ZH046/T7

AL050302.2, AL163203.2, AL049911.2, NM_014915.1, AB028997.1, AF090187.1, AL009051.1, AB011137.2, AL138654.1, AP000365.1, AP000548.1, AC011661.5, AP001302.1, AC004936.2, AC006157.2, Z99291.1, AL031599.1, 55 U40160.1, AC008526.5, AF156143.1, AC009402.3, AC004142.1, AC009513.2, AC006475.3, AC005760.1, AC005358.1, AL161595.2, Z11874.1, X70810.1, AL022605.3, U58744.1, X68658.1, X17051.1, AJ002397.1, AW373574.1, AW170035.1, AA759177.1, AL046701.1, AI957948.1, C87958.1, AW628933.1, AW469178.1, AW338178.1, AW331138.1, AW320227.1, AW171900.1, AW087179.1, AW052899.1, AW042526.1, AI972424.1, AI950371.1, AI920706.1, AI784583.1, AU072482.1, AI684965.1, AI460172.1, AI148480.1, AI093327.1, AI025802.1, AA970354.1, 60 AA708873.1, AA151117.1, AA150449.1, AA149652.1, AA136980.1, AA101607.1, AA071350.1, W73028.1, W35448.1, H97559.1, H96023.1, AL157387.2, AL162272.4, AC022596.4, AC015940.2, AC008088.2, AL354819.2, AL157695.2, AL138965.3, AC023067.3, AC016739.2, AC025384.2, AL161912.3, AC011983.3, AL137219.1, AL049185.4, AC040973.2, AC068690.1, AC026427.2, AC020901.5, AC010248.4, AC016684.1, AC026271.3, AC025076.3, AC051644.2, AC008390.6, AC011434.2, AC026081.2, AC008426.2, AC026395.2, AC016215.4, AC019042.3,

-234-

AA765215.1, AA745684.1, AA722275.1, AA720970.1, AA662367.1, AA593706.1, AA526463.1, AA525457.1,
 AA525257.1, AA447296.1, AA424313.1, AA404343.1, AA227150.1, AA149254.1, W61377.1, W23934.1, N94485.1,
 AW829301.1, AW646575.1, AW640539.1, AW516477.1, AW309883.1, AW139182.1, AI309747.1, AI301308.1,
 AI264756.1, AI088990.1, AI018656.1, AA830992.1, AA740946.1, AA703326.1, AA387253.1, AA074666.1, AL157387.2,
 5 AC027800.2, AP001004.2, AC022028.2, AC023528.3, AC024170.1, AL137244.14, AC010738.3, AC009406.3,
 AC015808.3, AC015945.3, AP001025.2, AC064867.1, AC011962.2, AC010832.3, AC022572.3, AC022551.3,
 AL157368.2, AC024583.3, AC067949.1, AC021029.3, AC007222.1, AL356137.2, AL163540.3, AL162395.2,
 AL137218.1, AL162272.4, AL138812.2, AL353620.1, AC021078.2, AC008818.4, AC011274.3, AC021738.2,
 AC009569.2, Z93245.1, AC011302.2, AC016720.4, AL139278.2, AC036237.4, AC009583.3, AC022952.3, AC006281.6,
 10 AC016466.3, AC021668.1, AL139008.8, AL137066.5, AL161629.2, Z95311.10, AC018922.8, AC018920.5, AC039057.3,
 AC020672.3, AC007670.2, AC013772.3, AC021801.1

SEQ ID NO: 557
 ZH1336/T3

15 AL163203.2, AL050302.2, AL049911.2, AC004802.1, AC002412.1, D88270.2, AC005856.1, U63630.1, AL035659.22,
 AL008635.1, AC016678.4, AC004019.20, AC024076.4, AC012078.3, AC009415.2, AC010072.5, AC005775.1,
 AC002400.1, AC000355.1, AF111170.3, AF225900.1, AF225899.1, AC005250.1, AC004008.1, AF176815.1,
 AC006014.2, AC006019.2, AC007274.2, AC007877.3, AC006368.2, AC007387.3, AC009516.19, AC002307.1,
 AC007214.13, AC007919.18, AC004814.2, AC007237.3, AC005089.2, AC007308.13, AC005488.2, AF134726.1,
 20 AC006241.1, AC006274.1, AC006130.1, AC004047.1, AC005722.1, L78833.1, AC004055.1, AC005523.1, AC005493.1,
 AC003682.1, AC004754.1, AC004449.1, AC004030.1, AC002558.1, AC002037.1, AL049759.10, AL035079.14,
 AL022163.1, AL021393.1, AL009051.1, AL023494.12, AL031602.14, AL163226.2, AL163222.2, AL160237.2,
 AL135940.11, AL137918.2, AL133238.2, AL035072.16, AL121601.13, AL034547.11, Z97630.11, Z68162.1, Z83001.1,
 Z99128.1, AL035604.15, Z84718.2, AL031657.2, AL009029.1, Z86061.1, AL033525.10, Z69706.1, Z69918.1, D87675.1,
 25 D00591.1, AP001681.1, AP001677.1, AP000365.1, AP000502.1, AP000957.2, AP000964.2, AP000307.1, AP000351.3,
 AP000548.1, AP000088.1, AC006581.16, AC005369.1, AL035695.17, AC006157.2, AW170035.1, AI375710.1,
 AA100715.1, N76504.1, AW812789.1, AW577142.1, AW026629.1, AI291588.1, AA601355.1, N98802.1, N78038.1,
 R92359.1, AW840742.1, AW627425.1, AW517377.1, AW503014.1, AW500125.1, AW298643.1, AW193265.1,
 AW103981.1, AW050726.1, AI990487.1, A1978583.1, AI950451.1, AI937125.1, AI907878.1, AL036070.1, AI801591.1,
 30 AI754658.1, AI687343.1, AI311927.1, AI133164.1, AA992337.1, AA834843.1, AA458863.1, AA362698.1, AA338904.1,
 AA314877.1, AI580781.1, AI168185.1, AI129968.1, AA788990.1, AA503577.1, AI283736.1, AI126468.1, AA621278.1,
 AA514328.1, W56548.1, H89487.1, AW298006.1, AI374954.1, AA639250.1, AA635049.1, AA192366.1, W27027.1,
 R07361.1, AW504900.1, AI870531.1, AL046701.1, AL043718.1, AI797903.1, AA339361.1, AW197070.1, AI652536.1,
 AI268666.1, AI126656.1, AA644357.1, AA280749.1, AA157017.1, AA031654.1, T23989.1, AI554395.1, AA989349.1,
 35 AA877743.1, AA831132.1, AA506458.1, AA447099.1, AW081941.1, AI808248.1, AA828419.1, AA487368.1, H39920.1,
 R72342.1, AW274289.1, AW005902.1, AI962478.1, AI351528.1, AI348589.1, AI198588.1, AI085314.1, AA947454.1,
 AA866064.1, AA832188.1, AA433996.1, AA284871.1, NS2189.1, H47518.1, AW271235.1, AL042073.1, AI244503.1,
 AL157387.2, AL162272.4, AC036208.2, AC023766.2, AC037459.1, AC021108.3, AC015559.1, AL157764.2,
 AC012300.2, AL353998.3, AL162261.2, AC067955.4, AC058791.2, AC020636.4, AC021103.6, AC060822.2,
 40 AC068889.4, AC020604.4, AC067948.3, AC061970.2, AC060764.2, AC044836.2, AC025643.3, AC006534.3,
 AC068720.1, AC021193.3, AC016739.2, AC034204.3, AC020900.3, AC011501.5, AC011452.5, AC008622.4,
 AC008403.5, AC008392.5, AC010354.4, AC010453.3, AC008839.4, AC032035.2, AC011511.4, AC022954.3,
 AC068034.1, AC067979.1, AC012439.4, AC026929.2, AC066599.1, AC063976.1, AC022311.4, AC009701.3,
 AC026839.2, AC023134.4, AC021674.3, AC023437.2, AC015992.3, AC015972.3, AC027119.2, AC025956.2,
 45 AC011473.3, AC023924.2, AC012118.2, AC025372.2, AC025371.2, AC022795.3, AC013331.4, AC024322.2,
 AC022489.3, AC021132.4, AC017063.5, AC023069.2, AC021287.4, AC010720.2, AC019238.2, AC023287.3,
 AC012356.3, AC009963.4, AC021922.1, AC009878.3, AC018408.1, AC018431.1, AC013560.2, AL355517.3,
 AL353648.4, AL049537.36, AL356354.1, AL139010.6, AL355586.4, AL355864.2, AL354683.2, AL353791.2,
 AL159152.6, AL355132.3, AL355922.1, AL138681.4, AL354867.1, AL161788.4, Z82207.1, AP001405.1, AP000729.1,
 50 AL356218.1, AC017099.3, AL356585.1,

SEQ ID NO: 558
 ZH1363/T3

AL050302.2, AL163203.2, AL049911.2, NM_014915.1, AB028997.1, AB011137.2, AL163202.2, AP001464.1,
 55 AP000542.1, AL031116.1, AL009051.1, AC007094.3, AC007970.3, AC018760.4, NM_003299.1, AC004936.2, U40939.1,
 AL161957.1, AL110120.11, AL078614.2, L42522.1, AP001465.1, X15187.1, AK000712.1, AC003096.2, AL133475.14,
 AL035690.10, AL035446.4, X03704.1, AW373574.1, AW614036.1, AW297165.1, AA521089.1, AA296607.1,
 AI988938.1, AI920892.1, AJ236531.1, AI180434.1, AI180260.1, AI176501.1, AA892166.1, AI058351.1, AI044283.1,
 AA533501.1, AW656081.1, AW656019.1, AW247016.1, AI984814.1, AI904680.1, AL046634.1, AL036452.1,
 60 AI770175.1, AI690374.1, AI446107.1, AA773607.1, AA398033.1, AA056391.1, W44909.1, W26464.1, R62950.1,
 M61978.1, AW662136.1, AW614443.1, AW613150.1, AW574899.1, AW574796.1, AW418540.1, AW418527.1,
 AW410433.1, AJ281607.1, AW320227.1, AW275258.1, AW244097.1, AV271334.1, AW088936.1, AW072295.1,
 AW005236.1, AI983391.1, AI924665.1, AV131385.1, AV081079.1, AI538762.1, AI368143.1, AI311542.1, AI302518.1,
 AI289573.1, AI200709.1, AI192436.1, AI069596.1, AI057937.1, AI008640.1, AA910838.1, AA856548.1, AA853969.1,

-245-

- AI585542.1, AA546260.1, AA263149.1, AI505847.1, AW106399.1, AA127538.1, AA285232.1, AW456026.1,
 AI588808.1, AI384994.1, AA981002.1, AW566712.1, AI787751.1, AW536727.1, AW413150.1, AV265194.1,
 AV219084.1, AI844907.1, AI842969.1, AV160844.1, AV165707.1, AV124038.1, AI779552.1, AV085555.1, AV047038.2,
 AV046630.2, AI325552.1, AU045190.1, AU018790.1, AU016981.1, AU016595.1, AU016513.1, AU015043.1,
 5 AU014858.1, AA252091.1, AA197255.1, AA108210.1, Z74637.1, F01019.1, AW705280.1, AJ281093.1, AW398070.1,
 AW398039.1, AW397497.1, AW397370.1, AW397141.1, AW397013.1, AW396970.1, AW396868.1, AW395825.1,
 AW395703.1, AW395679.1, AW395606.1, AW395515.1, AW225544.1, AW186387.1, AI973567.1, AI960869.1,
 AI941243.1, AI941225.1, AI940836.1, AI883261.1, AL079496.1, AI795023.1, AI759696.1, AI748161.1, AI748087.1,
 AI736054.1, AI736025.1, AI736012.1, AI406060.1, AI109316.1, AI109205.1, AI063325.1, AA902197.1, AA784297.1,
 10 AA553106.1, AA497210.1, AA466795.1, AA462438.1, AA432643.1, AA313904.1, C13433.1, F12959.1, AV185121.1,
 C66585.1, AL138878.4, AL139095.5, AC010893.4, AL109933.21, AC025343.2, AC008760.4, AC024364.3, AC063933.3,
 AC025925.2, AC021185.2, AC023350.1, AC016546.4, AC034197.2, AC021814.2, AC026219.1, AC016723.4,
 AC007218.2, AC012373.13, AC022890.1, AC017288.1, AC018551.1, AC015313.1, AL139011.6, AP001075.2,
 AC025685.2, AC061978.2, AC008427.5, AC021869.6, AC027631.2, AC026233.2, AC013364.7, AC013350.6,
 15 AC026232.1, AC018740.2, AC023364.3, AC008050.3, AC024301.1, AC010003.5, AC009368.5, AC017158.1,
 AC008367.3, AC020124.1, AF215848.1, AC008236.3, AC017830.1, AC017944.1, AC012952.1, AL122026.2,
 AL049185.4, AC068808.4, AC037424.7, AC016634.4, AC008904.3, AC009550.3, AC021972.2, AC020804.2,
 AC021305.3, AC023103.3, AC020080.1, AC020324.1, AC012195.2, AL138711.3, AL160153.4, AL031745.7
 20 SEQ ID NO: 593
 ZH204/T7
 AC012599.8, AC004092.1, L78822.1, L04666.1, NC_001146.1, NM_005711.1, AC005919.1, AC005788.1, AC003036.1,
 U70312.1, AF003530.1, X74595.1, Z71448.1, L20973.1, L19930.1, AI686567.1, AW073551.1, AA007617.1, AA702832.1,
 AA778768.1, AA127539.1, AA085379.1, F31106.1, AW196506.1, F36537.1, AW137246.1, AW268860.1, AW582844.1,
 25 AI118179.1, AI651413.1, AW324433.1, AI465698.1, AA073164.1, AW390105.1, AA856137.1, AA577233.1,
 AA648320.1, AI990395.1, AA072738.1, AI904456.1, AU024036.1, AI702678.1, AA127538.1, AI904448.1, AV359288.1,
 AI420526.1, AI221321.1, AV292110.1, AI616122.1, AA693126.1, AW215056.1, AV318953.1, AI561593.1, AA153299.1,
 AA007643.1, AA689696.1, AW431906.1, AV374296.1, AV155600.1, AL138878.4, AL139095.5, AC005842.6,
 AC024410.2, AC053543.3, AC008502.4, AC024218.2, AC009292.7, AC055730.3, AC009362.6, AC007351.16,
 30 AC055710.3, AC025577.10, AC024219.7, AC024146.5, AC022265.2, AC068656.1, AC016639.5, AC008422.4,
 AC016632.4, AC008914.3, AC025763.2, AC024164.2, AC023194.3, AC034249.1, AC016441.4, AC024469.1,
 AL049185.4

35

Table 2: Relation between nucleotide sequences and polypeptide sequences

Nucleic acid SEQ ID NO	Polypeptide SEQ ID NO	Nucleic acid SEQ ID NO	Polypeptide SEQ ID NO	Nucleic acid SEQ ID NO	Polypeptide SEQ ID NO
1	665	199	818	397	1117
2	678	200	819, 820	398	—
3	679	201	821	399	1118
4	761, 762	202	822	400	1119
5	763, 764, 765	203	823, 824	401	1120
6	782	204	825, 826	402	--
7	783	205	827, 828, 829	403	1121
8	767	206	830, 831	404	1122
9	604	207	832	405	1123
10	--	208	833	406	1124
11	606	209	834, 835	407	--
12	624	210	836, 837	408	--
13	599	211	838	409	1125
14	776, 777, 778, 779	212	839, 840, 841	410	--

-246-

15	780, 781	213	842, 843	411	1126
16	802	214	844, 845, 846	412	1127
17	803	215	847, 848	413	1128
18	607	216	849	414	-
19	594	217	850, 851, 852	415	1129
20	595, 596, 597	218	853, 854	416	1130, 1131, 1132
21	--	219	--	417	1133
22	598	220	--	418	1134
23		221	855, 856, 857	419	1135, 1136
24	600	222	858, 859, 860	420	1137, 1138
25		223	861, 862, 863	421	1139
26	601	224	864	422	--
27	-	225	865	423	1140
28	602	226	866, 867, 868	424	1141, 1142
29	603	227	869, 870	425	1143
30	605	228	871	426	1144, 1145
31	--	229	872	427	--
32	608	230	873	428	1146, 1147
33	609	231	874	429	1148, 1149, 1150
34	610	232	875, 876, 877	430	--
35	611, 612	233	878	431	1151
36	-	234	879, 880	432	1152
37	-	235	881, 882	433	--
38	613	236	883	434	1153
39	614	237	884, 885	435	1154
40	-	238	886	436	--
41	615	239	887, 888	437	1155
42	-	240	--	438	1156
43	616	241	889, 890	439	--
44	--	242	891, 892	440	1157
45	617, 618	243	893, 894	441	1158
46	619	244	895	442	1159
47	620	245	896	443	1160
48	621	246	897	444	1161
49	-	247	898	445	1162
50	622	248	899	446	--
51	623	249	900, 901	447	1163
52	625	250	902	448	--
53	--	251	903	449	1164
54	626	252	904	450	1165
55	627, 628	253	905	451	1166

-247-

56	629	254	906	452	--
57	--	255	907, 908, 909	453	1167
58	630	256	910, 911	454	--
59	--	257	912, 913, 914	455	1168
60	631	258	915	456	--
61	632	259	916	457	1169
62	--	260	917, 918, 918	458	1170
63	633	261	920, 921	459	1171
64	634	262	922, 923	460	1172
65	635	263	924, 925, 926	461	1173, 1174, 1175
66	--	264	927, 928, 929	462	1176, 1177
67	636	265	930, 931	463	1178
68	--	266	932	464	1179
69	637	267	933, 934	465	1180
70	--	268	935	466	1181, 1182, 1183
71	--	269	936	467	1184
72	--	270	937, 938	468	1185
73	--	271	939	469	1186
74	--	272	940	470	1187
75	--	273	941, 942	471	1188
76	--	274	943, 944	472	--
77	638	275	945	473	1189
78	--	276	946, 947, 948	474	1190
79	--	277	949	475	--
80	--	278	950, 951, 952	476	1191, 1192
81	639	279	953, 954	477	1193
82	--	280	955, 956	478	1194
83	640	281	--	479	1195
84	641	282	957	480	1196
85	--	283	958	481	1197
86	642	284	959	482	1198
87	--	285	960, 961	483	1199
88	643	286	--	484	1200, 1201
89	--	287	962, 963, 964	485	1202
90	644	288	965	486	1203
91	--	289	966	487	1204
92	645	290	967, 968, 969, 970	488	1205
93	646	291	971, 972, 973, 974	489	1206
94	647	292	975, 976	490	1207
95	648	293	977, 978, 979,	491	1208

-248-

			980		
96	649, 650	294	981, 982, 983	492	1209
97	651, 652	295	984, 985	493	1210
98	653	296	986, 987	494	--
99	654	297	988, 989	495	1211
100	655	298	990	496	1212
101	656	299	991	497	1213
102	657	300	992	498	--
103	658	301	993	499	1214
104	659, 660	302	994	500	--
105	661	303	995	501	1215
106	662, 663	304	996	502	1216
107	664	305	997	503	1217
108	666	306	998	504	1218
109	667	307	999	505	1219
110	--	308	1000, 1001	506	1220
111	668, 669, 670	309	1002, 1003	507	1221
112	671, 672, 673, 674, 675	310	1004, 1005	508	1222
113	676	311	1006, 1007, 1008	509	--
114	677	312	1009, 1010	510	1223
115	680	313	1011, 1012, 1013	511	1224
116	681	314	1014, 1015	512	1225
117	682, 683	315	1016, 1017	513	1226
118	684	316	1018, 1019	514	1227
119	685	317	1020	515	1228
120	686	318	1021	516	1229, 1230
121	687, 688	319	1022	517	1231
122	689, 690, 691	320	1023	518	--
123	692	321	1024, 1025, 1026	519	1232
124	693, 694	322	1027, 1028, 1029, 1030	520	1233
125	695	323	1031, 1032, 1033	521	1234
126	696	324	1034, 1035, 1036	522	1235
127	697	325	1037, 1038	523	--
128	698	326	1039, 1040	524	1236
129	699	327	1041, 1042, 1043, 1044	525	--
130	700, 701, 702	328	1045, 1046	526	1237
131	703, 704	329	1047	527	--
132	705	330	1048, 1049	528	1238
133	706, 707	331	1050	529	1239

-249-

134	708	332	1051, 1052	530	1240, 1241
135	709, 710	333	1053	531	1242, 1243, 1244
136	711	334	1054, 1055	532	1245
137	712	335	1056	533	1246, 1247
138	713	336	1057, 1058	534	--
139	714	337	1059	535	1248
140	715	338	1060, 1061	536	1249, 1250, 1251
141	716	339	1062	537	1252, 1253, 1254
142	717	340	1063	538	--
143	718	341	1064, 1065	539	1255
144	719	342	1066	540	1256
145	720	343	1067, 1068	541	1257
146	721	344	1069	542	1258
147	722	345	1070, 1071	543	1259, 1260
148	723, 724	346	1072	544	1261
149	725, 726	347	1073, 1074	545	1262
150	727	348	1075, 1076	546	1263
151	728, 729	349	1077, 1078, 1079	547	1264
152	730, 731	350	1080	548	1265, 1266
153	732, 733	351	1081	549	1267
154	734, 735, 736	352	1082, 1083	550	--
155	737	353	1084	551	1268
156	738	354	1085	552	1269, 1270
157	739, 740	355	--	553	1271, 1272
158	741	356	1086	554	1273, 1274
159	742, 743, 744	357	--	555	1275
160	745	358	1087	556	1276
161	746	359	1088	557	1277
162	747, 748	360	1089	558	1278
163	749	361	1090	559	1279
164	750	362	1091	560	1280
165	751, 752	363	--	561	1281
166	753	364	1092	562	1282, 1283
167	754, 755, 756	365	--	563	1284, 1285
168	757, 758, 759, 760	366	1093	564	1286, 1287
169	766	367	1094	565	1288, 1289, 1290
170	768	368	1095	566	1291, 1292
171	769	369	1096	567	1293, 1294
172	770	370	1097	568	1295
173	771	371	--	569	1296, 1297, 1298
174	772	372	--	570	1299

-250-

175	773, 774, 775	373	1098	571	1300
176	784	374	1099	572	1301
177	785	375	--	573	1302, 1303
178	786, 787, 788	376	1100	574	1304
179	--	377	1101	575	--
180	789	378	1102	576	1305, 1306
181	790	379	1103	577	1307
182	791	380	1104	578	1308
183	792, 793	381	1105	579	1309
184	794	382	1106	580	1310
185	795, 796	383	1107	581	1311, 1312, 1313
186	797, 798, 799, 800	384	1108	582	1314, 1315, 1316
187	801	385	--	583	1317, 1318
188	804, 805	386	1109	584	1319
189	806	387	1110	585	1320, 1321
190	807	388	1111	586	1322
191	808, 809	389	--	587	1323, 1324
192	810	390	1112	588	1325
193	811	391	--	589	1326
194	812	392	1113	590	1327
195	813	393	1114	591	1328, 1329
196	814	394	1115	592	1330, 1331
197	815	395	--	593	1332
198	816, 817	396	1116		

Example 2: Preparation of recombinant cancer associated antigens

To facilitate screening of patients' sera for antibodies reactive with cancer associated antigens, for example by ELISA, recombinant proteins are prepared according to standard procedures. Where gaps exist in the gene sequences represented by the clones disclosed herein, or where flanking sequences are desired, such nucleic acid sequences can be isolated according to standard procedures. For example, where 5' and 3' clones of a gene sequence are known, PCR primers can be designed for amplification of the nucleotide sequence between the clones. Flanking sequences can be isolated using procedures such as RACE PCR. Such sequences also can be isolated by standard hybridization cloning protocols.

In one method of preparing recombinant cancer associated antigens, the clones encoding cancer associated antigens are subcloned into a baculovirus expression vector, and the recombinant expression vectors are introduced into appropriate insect cells.

-251-

Baculovirus/insect cloning systems are preferred because post-translational modifications are carried out in the insect cells. Another preferred eukaryotic system is the *Drosophila* Expression System from Invitrogen. Clones which express high amounts of the recombinant protein are selected and used to produce the recombinant proteins. The recombinant proteins 5 are tested for antibody recognition using serum from the patient which was used to isolated the particular clone, or in the case of cancer associated antigens recognized by allogeneic sera, by the sera from any of the patients used to isolate the clones or sera which recognize the clones' gene products.

Alternatively, the cancer associated antigen clones are inserted into a prokaryotic 10 expression vector for production of recombinant proteins in bacteria. Other systems, including yeast expression systems and mammalian cell culture systems also can be used.

Example 3: Preparation of antibodies to cancer associated antigens

The recombinant cancer associated antigens produced as in Example 2 above are used 15 to generate polyclonal antisera and monoclonal antibodies according to standard procedures. The antisera and antibodies so produced are tested for correct recognition of the cancer associated antigens by using the antisera/antibodies in assays of cell extracts of patients known to express the particular cancer associated antigen (e.g. an ELISA assay). These 20 antibodies can be used for experimental purposes (e.g. localization of the cancer associated antigens, immunoprecipitations, Western blots, etc.) as well as diagnostic purposes (e.g., testing extracts of tissue biopsies, testing for the presence of cancer associated antigens).

Example 4: Expression of breast, gastric and prostate cancer associated antigens in cancers of similar and different origin.

25 The expression of one or more of the breast, gastric and/or prostate cancer associated antigens is tested in a range of tumor samples to determine which, if any, other malignancies should be diagnosed and/or treated by the methods described herein. Tumor cell lines and tumor samples are tested for cancer associated antigen expression, preferably by RT-PCR according to standard procedures. Northern blots also are used to test the expression of the 30 cancer associated antigens. Antibody based assays, such as ELISA and western blot, also can be used to determine protein expression. A preferred method of testing expression of cancer associated antigens (in other cancers and in additional same type cancer patients) is allogeneic

-252-

serotyping using a modified SEREX protocol (as described above).

In all of the foregoing, extracts from the tumors of patients who provided sera for the initial isolation of the cancer associated antigens are used as positive controls. The cells containing recombinant expression vectors described in the Examples above also can be used
5 as positive controls.

The results generated from the foregoing experiments provide panels of multiple cancer associated nucleic acids and/or polypeptides for use in diagnostic (e.g. determining the existence of cancer, determining the prognosis of a patient undergoing therapy, etc.) and therapeutic methods (e.g., vaccine composition, etc.).

10

Example 5: HLA typing of patients positive for cancer associated antigen

To determine which HLA molecules present peptides derived from the cancer associated antigens, cells of the patients which express the breast and/or gastric cancer associated antigens are HLA typed. Peripheral blood lymphocytes are taken from the patient
15 and typed for HLA class I or class II, as well as for the particular subtype of class I or class II. Tumor biopsy samples also can be used for typing. HLA typing can be carried out by any of the standard methods in the art of clinical immunology, such as by recognition by specific monoclonal antibodies, or by HLA allele-specific PCR (e.g. as described in WO97/31126).

20 **Example 6: Characterization of cancer associated antigen peptides presented by MHC class I and class II molecules.**

Antigens which provoke an antibody response in a subject may also provoke a cell-mediated immune response. Cells process proteins into peptides for presentation on MHC class I or class II molecules on the cell surface for immune surveillance. Peptides presented
25 by certain MHC/HLA molecules generally conform to motifs. These motifs are known in some cases, and can be used to screen the breast and/or gastric cancer associated antigens for the presence of potential class I and/or class II peptides. Summaries of class I and class II motifs have been published (e.g., Rammensee et al., *Immunogenetics* 41:178-228, 1995). Based on the results of experiments such as those described above, the HLA types which
30 present the individual breast cancer associated antigens are known. Motifs of peptides presented by these HLA molecules thus are preferentially searched.

One also can search for class I and class II motifs using computer algorithms. For

example, computer programs for predicting potential CTL epitopes based on known class I motifs has been described (see, e.g., Parker et al., *J. Immunol.* 152:163, 1994; D'Amaro et al., *Human Immunol.* 43:13-18, 1995; Drijfhout et al., *Human Immunol.* 43:1-12, 1995). HLA binding predictions can conveniently be made using an algorithm available via the Internet on the National Institutes of Health World Wide Web site at URL <http://bimas.dcrt.nih.gov>. Methods for determining HLA class II peptides and making substitutions thereto are also known (see, e.g. International applications PCT/US96/03182 and PCT/US98/01373). Computer software for selecting HLA class II binding peptides is also available (TEPITOPE; Sturniolo et al., *Nature Biotechnol.* 17:555-561, 1999; Manici et al., *J. Exp. Med.* 189:871-876, 1999). Peptides which are thus selected can be for inducing specific CD4⁺ lymphocytes and identification of peptides. Additional methods of selecting and testing peptides for HLA class II binding are well known in the art.

Example 7: Identification of the portion of a cancer associated polypeptide encoding an antigen

To determine if the cancer associated antigens isolated as described above can provoke a cytolytic T lymphocyte response, the following method is performed. CTL clones are generated by stimulating the peripheral blood lymphocytes (PBLs) of a patient with autologous normal cells transfected with one of the clones encoding a cancer associated antigen polypeptide or with irradiated PBLs loaded with synthetic peptides corresponding to the putative protein and matching the consensus for the appropriate HLA class I molecule (as described above) to localize an antigenic peptide within the cancer associated antigen clone (see, e.g., Knuth et al., *Proc. Natl. Acad. Sci. USA* 81:3511-3515, 1984; van der Bruggen et al., *Eur. J. Immunol.* 24:3038-3043, 1994). These CTL clones are screened for specificity against COS cells transfected with the cancer associated antigen clone and autologous HLA alleles as described by Brichard et al. (*Eur. J. Immunol.* 26:224-230, 1996). CTL recognition of a cancer associated antigen is determined by measuring release of TNF from the cytolytic T lymphocyte or by ⁵¹Cr release assay (Herin et al., *Int. J. Cancer* 39:390-396, 1987). If a CTL clone specifically recognizes a transfected COS cell, then shorter fragments of the cancer associated antigen clone transfected in that COS cell are tested to identify the region of the gene that encodes the peptide. Fragments of the cancer associated antigen clone are prepared by exonuclease III digestion or other standard molecular biology methods. Synthetic peptides

are prepared to confirm the exact sequence of the antigen.

Optionally, shorter fragments of cancer associated antigen cDNAs are generated by PCR. Shorter fragments are used to provoke TNF release or ^{51}Cr release as above.

Synthetic peptides corresponding to portions of the shortest fragment of the cancer associated antigen clone which provokes TNF release are prepared. Progressively shorter peptides are synthesized to determine the optimal cancer associated antigen tumor rejection antigen peptides for a given HLA molecule.

A similar method is performed to determine if the cancer associated antigen contains one or more HLA class II peptides recognized by T cells. One can search the sequence of the cancer associated antigen polypeptides for HLA class II motifs as described above. In contrast to class I peptides, class II peptides are presented by a limited number of cell types. Thus for these experiments, dendritic cells or B cell clones which express HLA class II molecules preferably are used.

15 Example 8: Recognition of cancer antigens by cancer patient sera

Several of the cancer antigen identified herein were tested for reactivity with sera from normal and breast cancer patients according to standard procedures (e.g., the SEREX procedure outlined above).

20 Table 3: Serology of antigens

SEQ ID NO	Gene/Clone	Breast Cancer Patient Sera	Normal Sera
1	Br-38/HSP105 (MK)	6/31	0/30
2,3	Br-39/HSP105 (MK)	3/31	0/30
4,5	RGS-GAIP interacting protein GIPC (MK)	3/31	0/30
6,7	NS1-binding protein/KIAA0850 (MK)	3/31	0/30
8	Opa-interacting protein OIP2 (MK)	3/31	0/30
9,10	Kinesin family protein 3B (KIF3B) (MT)	2/31	0/30
11	Endothelial-monocyte activating protein (EMAP2) (MT)	2/31	0/30
12	Unknown TOM1 protein (MT311)	2/31	0/30
13	Outer mitochondrial membrane protein 34kDa (MT)	1/31	0/30

-255-

14,15	IPL (MK)	1/31	0/30
16,17	Mus ACF7 neural isoform (MK)	1/31	0/30
18	Cyclin D3 (MT)	1/31	0/30

The data show that proteins encoded by SEQ ID NO:1-12 were recognized by multiple breast cancer patients' sera, but not by control individuals' sera. Proteins encoded by SEQ ID NO:13-18 were recognized by only a single breast cancer patient's sera, but not by control

5 individuals' sera. The

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention
10 described herein. Such equivalents are intended to be encompassed by the following claims.

All references disclosed herein are incorporated by reference in their entirety.

We claim:

-256-

Claims

1. A method of diagnosing a disorder characterized by expression of a human cancer associated antigen precursor coded for by a nucleic acid molecule, comprising:
 - 5 contacting a biological sample isolated from a subject with an agent that specifically binds to the nucleic acid molecule, an expression product thereof, or a fragment of an expression product thereof complexed with an HLA molecule, wherein the nucleic acid molecule is a NA Group 1 nucleic acid molecule, and
 - 10 determining the interaction between the agent and the nucleic acid molecule or the expression product as a determination of the disorder.
- 15 2. The method of claim 1, wherein the agent is selected from the group consisting of
 - (a) a nucleic acid molecule comprising NA group 1 nucleic acid molecules or a fragment thereof,
 - (b) a nucleic acid molecule comprising NA group 3 nucleic acid molecules or a fragment thereof,
 - (c) a nucleic acid molecule comprising NA group 5 nucleic acid molecules or a fragment thereof,
 - (d) an antibody that binds to an expression product of NA group 1 nucleic acids,
 - (e) an antibody that binds to an expression product of NA group 3 nucleic acids,
 - 20 (f) an antibody that binds to an expression product of NA group 5 nucleic acids,
 - (g) an agent that binds to a complex of an HLA molecule and a fragment of an expression product of a NA group 1 nucleic acid,
 - (h) an agent that binds to a complex of an HLA molecule and a fragment of an expression product of a NA group 3 nucleic acid, and
 - 25 (i) an agent that binds to a complex of an HLA molecule and a fragment of an expression product of a NA group 5 nucleic acid.
- 30 3. The method of claim 1, wherein the disorder is characterized by expression of a plurality of human cancer associated antigen precursors and wherein the agent is a plurality of agents, each of which is specific for a different human cancer associated antigen precursor, and wherein said plurality of agents is at least 2, at least 3, at least 4, at least 4, at least 6, at least 7, or at least 8, at least 9 or at least 10 such agents.

4. The method of claims 1-3, wherein the agent is specific for a human cancer associated antigen precursor that is a breast cancer associated antigen precursor.

5 5. A method for determining regression, progression or onset of a condition characterized by expression of abnormal levels of a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule, comprising

monitoring a sample, from a patient who has or is suspected of having the condition, for a parameter selected from the group consisting of

10 (i) the protein,
(ii) a peptide derived from the protein,
(iii) an antibody which selectively binds the protein or peptide, and
(iv) cytolytic T cells specific for a complex of the peptide derived from the protein and an MHC molecule,

15 as a determination of regression, progression or onset of said condition.

6. The method of claim 5, wherein the sample is a body fluid, a body effusion or a tissue.

7. The method of claim 5, wherein the step of monitoring comprises contacting the 20 sample with a detectable agent selected from the group consisting of

(a) an antibody which selectively binds the protein of (i), or the peptide of (ii),
(b) a protein or peptide which binds the antibody of (iii), and
(c) a cell which presents the complex of the peptide and MHC molecule of (iv).

25 8. The method of claim 7, wherein the antibody, the protein, the peptide or the cell is labeled with a radioactive label or an enzyme.

9. The method of claim 5, comprising assaying the sample for the peptide.

30 10. The method of claim 5, wherein the nucleic acid molecule is a NA Group 3 molecule.

11. The method of claim 5, wherein the nucleic acid molecule is a NA Group 5 molecule.

12. The method of claim 5, wherein the protein is a plurality of proteins, the parameter is a plurality of parameters, each of the plurality of parameters being specific for a different one of the plurality of proteins, at least one of which is a cancer associated protein encoded by a NA
5 Group 1 molecule.
13. A pharmaceutical preparation for a human subject comprising
an agent which when administered to the subject enriches selectively the presence of complexes of an HLA molecule and a human cancer associated antigen, and
10 a pharmaceutically acceptable carrier, wherein the human cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule comprises a NA Group 1 molecule.
14. The pharmaceutical preparation of claim 13, wherein the agent comprises a plurality of agents, each of which enriches selectively in the subject complexes of an HLA molecule and a different human cancer associated antigen, wherein at least one of the human cancer associated antigens is encoded by a NA Group 1 molecule.
15
15. The pharmaceutical preparation of claim 14, wherein the plurality is at least two, at
20 least three, at least four or at least 5 different such agents.
16. The pharmaceutical preparation of claim 13, wherein the nucleic acid molecule is a NA Group 3 nucleic acid molecule.
25
17. The pharmaceutical preparation of claim 13, wherein the agent is selected from the group consisting of
 (1) an isolated polypeptide comprising the human cancer associated antigen, or a functional variant thereof,
 (2) an isolated nucleic acid operably linked to a promoter for expressing the isolated
30 polypeptide, or functional variant thereof,
 (3) a host cell expressing the isolated polypeptide, or functional variant thereof, and
 (4) isolated complexes of the polypeptide, or functional variants thereof, and an HLA

molecule.

18. The pharmaceutical preparation of claims 13-17, further comprising an adjuvant.
- 5 19. The pharmaceutical preparation of claim 13, wherein the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof, and wherein the cell is nonproliferative.
- 10 20. The pharmaceutical preparation of claim 13, wherein the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof, and wherein the cell expresses an HLA molecule that binds the polypeptide.
- 15 21. The pharmaceutical preparation of claim 13, wherein the agent is at least two, at least three, at least four or at least five different polypeptides, each coding for a different human cancer associated antigen or functional variant thereof, wherein at least one of the human cancer associated antigens is encoded by a NA Group 1 molecule.
- 20 22. The pharmaceutical preparation of claim 13, wherein the agent is a PP Group 2 polypeptide.
23. The pharmaceutical preparation of claim 13, wherein the agent is a PP Group 3 polypeptide or a PP Group 4 polypeptide.
- 25 24. The pharmaceutical preparation of claim 20, wherein the cell expresses one or both of the polypeptide and HLA molecule recombinantly.
25. The pharmaceutical preparation of claim 20, wherein the cell is nonproliferative.
- 30 26. A composition comprising
an isolated agent that binds selectively a PP Group 1 polypeptide.
27. The composition of matter of claim 26, wherein the agent binds selectively a PP Group

2 polypeptide.

28. The composition of matter of claim 26, wherein the agent binds selectively a PP Group
3 polypeptide.

5

29. The composition of matter of claim 26, wherein the agent binds selectively a PP Group
4 polypeptide.

30. The composition of matter of claim 26, wherein the agent binds selectively a PP Group
10 5 polypeptide.

31. The composition of claims 26-30, wherein the agent is a plurality of different agents
that bind selectively at least two, at least three, at least four, or at least five different such
polypeptides.

15

32. The composition of claims 26-30, wherein the agent is an antibody.

33. The composition of claim 31, wherein the agent is an antibody.

20 34. A composition of matter comprising
a conjugate of the agent of claims 26-30 and a therapeutic or diagnostic agent.

35. A composition of matter comprising
a conjugate of the agent of claim 31 and a therapeutic or diagnostic agent.

25

36. The composition of matter of claim 34, wherein the conjugate is of the agent and a
therapeutic or diagnostic that is a toxin.

30 37. A pharmaceutical composition comprising an isolated nucleic acid molecule selected
from the group consisting of NA Group 1 molecules and NA Group 2 molecules, and a
pharmaceutically acceptable carrier.

-261-

38. The pharmaceutical composition of claim 37, wherein the isolated nucleic acid molecule comprises a NA Group 3 or NA Group 4 molecule.

39. The pharmaceutical composition of claim 37, wherein the isolated nucleic acid 5 molecule comprises at least two isolated nucleic acid molecules coding for two different polypeptides, each polypeptide comprising a different human cancer associated antigen.

40. The pharmaceutical composition of claims 37-39 further comprising an expression vector with a promoter operably linked to the isolated nucleic acid molecule.

10

41. The pharmaceutical composition of claims 37-39 further comprising a host cell recombinantly expressing the isolated nucleic acid molecule.

15

42. A pharmaceutical composition comprising an isolated polypeptide comprising a PP Group 1 or a PP Group 2 polypeptide, and a pharmaceutically acceptable carrier.

43. The pharmaceutical composition of claim 42, wherein the isolated polypeptide comprises a PP Group 3 or a PP Group 4 polypeptide.

20

44. The pharmaceutical composition of claim 42, wherein the isolated polypeptide comprises at least two different polypeptides, each comprising a different human cancer associated antigen.

25

45. The pharmaceutical composition of claim 42, wherein the isolated polypeptides are breast cancer polypeptides or HLA binding fragments thereof.

46. The pharmaceutical composition of claim 42, wherein the isolated polypeptides are 30 gastric cancer polypeptides or HLA binding fragments thereof.

30

47. The pharmaceutical composition of claims 42-46, further comprising an adjuvant.

-262-

48. An isolated nucleic acid molecule comprising a NA Group 3 molecule.

49. An isolated nucleic acid molecule comprising a NA Group 4 molecule.

5 50. An isolated nucleic acid molecule selected from the group consisting of
(a) a fragment of a nucleic acid molecule having a nucleotide sequence selected from
the group consisting of nucleotide sequences set forth as SEQ ID NOs:1-593, of sufficient
length to represent a sequence unique within the human genome, and identifying a nucleic
acid encoding a human cancer associated antigen precursor,
10 (b) complements of (a),
provided that the fragment includes a sequence of contiguous nucleotides which is not
identical to any sequence selected from the sequence group consisting of
(1) sequences having the GenBank accession numbers of Table 1, and other publicly
available sequences,
15 (2) complements of (1), and
(3) fragments of (1) and (2).

51. The isolated nucleic acid molecule of claim 50, wherein the sequence of contiguous
nucleotides is selected from the group consisting of:

20 (1) at least two contiguous nucleotides nonidentical to the sequence group,
(2) at least three contiguous nucleotides nonidentical to the sequence group,
(3) at least four contiguous nucleotides nonidentical to the sequence group,
(4) at least five contiguous nucleotides nonidentical to the sequence group,
(5) at least six contiguous nucleotides nonidentical to the sequence group,
25 (6) at least seven contiguous nucleotides nonidentical to the sequence group.

52. The isolated nucleic acid molecule of claim 50, wherein the fragment has a size
selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides,
14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24
30 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides,
100 nucleotides, and 200 nucleotides.

-263-

53. The isolated nucleic acid molecule of claim 50, wherein the molecule encodes a polypeptide which, or a fragment of which, binds a human HLA receptor or a human antibody.
- 5 54. An expression vector comprising an isolated nucleic acid molecule of any of claims 48-53 operably linked to a promoter.
- 10 55. An expression vector comprising a nucleic acid operably linked to a promoter, wherein the nucleic acid is a NA Group 2 molecule.
- 15 56. An expression vector comprising a NA Group 1 or Group 2 molecule and a nucleic acid encoding an HLA molecule.
57. A host cell transformed or transfected with an expression vector of claim 54.
- 15 58. A host cell transformed or transfected with an expression vector of claims 55 or 56.
59. A host cell transformed or transfected with an expression vector of claim 54 and further comprising a nucleic acid encoding HLA.
- 20 60. A host cell transformed or transfected with an expression vector of claim 55 and further comprising a nucleic acid encoding HLA.
- 25 61. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 48 or claim 49.
62. A fragment of the polypeptide of claim 61 which is immunogenic.
- 30 63. The fragment of claim 62, wherein the fragment, or a portion of the fragment, binds HLA or a human antibody.
64. An isolated fragment of a human cancer associated antigen precursor which, or portion

of which, binds HLA or a human antibody, wherein the precursor is encoded by a nucleic acid molecule that is a NA Group 1 molecule.

65. The fragment of claim 64, wherein the fragment is part of a complex with HLA.

5

66. The fragment of claim 65, wherein the fragment is between 8 and 12 amino acids in length.

10 67. An isolated polypeptide comprising a fragment of the polypeptide of claim 61 of sufficient length to represent a sequence unique within the human genome and identifying a polypeptide that is a human cancer associated antigen precursor.

68. A kit for detecting the presence of the expression of a human cancer associated antigen precursor comprising

15 a pair of isolated nucleic acid molecules each of which consists essentially of a molecule selected from the group consisting of (a) a 12-32 nucleotide contiguous segment of the nucleotide sequence of any of the NA Group 1 molecules and (b) complements of (a), wherein the contiguous segments are nonoverlapping.

20 69. The kit of claim 68, wherein the pair of isolated nucleic acid molecules is constructed and arranged to selectively amplify an isolated nucleic acid molecule that is a NA Group 3 molecule.

70. A method for treating a subject with a disorder characterized by expression of a human 25 cancer associated antigen precursor, comprising

administering to the subject an amount of an agent, which enriches selectively in the subject the presence of complexes of an HLA molecule and a human cancer associated antigen, effective to ameliorate the disorder, wherein the human cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule 30 selected from the group consisting of

- (a) a nucleic acid molecule comprising NA group 1 nucleic acid molecules,
- (b) a nucleic acid molecule comprising NA group 3 nucleic acid molecules,

-265-

(c) a nucleic acid molecule comprising NA group 5 nucleic acid molecules.

71. The method of claim 70, wherein the disorder is characterized by expression of a plurality of human cancer associated antigen precursors and wherein the agent is a plurality of agents, each of which enriches selectively in the subject the presence of complexes of an HLA molecule and a different human cancer associated antigen, wherein at least one of the human cancer associated antigens is encoded by a NA Group 1 molecule.
72. The method of claim 71, wherein the plurality is at least 2, at least 3, at least 4, or at least 5 such agents.
73. The method of claims 70-72, wherein the agent is an isolated polypeptide selected from the group consisting of PP Group 1, PP Group 2, PP Group 3, PP Group 4, and PP Group 5.
74. The method of claims 70-72, wherein the disorder is cancer.
75. The method of claims 73, wherein the disorder is cancer.
76. A method for treating a subject having a condition characterized by expression of a human cancer associated antigen precursor in cells of the subject, comprising:
- (i) removing an immunoreactive cell containing sample from the subject,
 - (ii) contacting the immunoreactive cell containing sample to the host cell under conditions favoring production of cytolytic T cells against a human cancer associated antigen which is a fragment of the precursor,
 - (iii) introducing the cytolytic T cells to the subject in an amount effective to lyse cells which express the human cancer associated antigen, wherein the host cell is transformed or transfected with an expression vector comprising an isolated nucleic acid molecule operably linked to a promoter, the isolated nucleic acid molecule being selected from the group of nucleic acid molecules consisting of NA Group 1, NA Group 2, NA Group 3, NA Group 4, and NA Group 5.

77. The method of claim 76, wherein the host cell recombinantly expresses an HLA molecule which binds the human cancer associated antigen.

78. The method of claim 76, wherein the host cell endogenously expresses an HLA molecule which binds the human cancer associated antigen.
5

79. A method for treating a subject having a condition characterized by expression of a human cancer associated antigen precursor in cells of the subject, comprising:

(i) identifying a nucleic acid molecule expressed by the cells associated with said
10 condition, wherein said nucleic acid molecule is a NA Group 1 molecule;

(ii) transfecting a host cell with a nucleic acid selected from the group consisting of (a) the nucleic acid molecule identified, (b) a fragment of the nucleic acid identified which includes a segment coding for a human cancer associated antigen, (c) deletions, substitutions or additions to (a) or (b), and (d) degenerates of (a), (b), or (c);
15

(iii) culturing said transfected host cells to express the transfected nucleic acid molecule, and;
20

(iv) introducing an amount of said host cells or an extract thereof to the subject effective to increase an immune response against the cells of the subject associated with the condition.

80. The method of claim 79, further comprising identifying an MHC molecule which presents a portion of an expression product of the nucleic acid molecule, wherein the host cell expresses the same MHC molecule as identified and wherein the host cell presents an MHC binding portion of the expression product of the nucleic acid molecule.
25

81. The method of claim 79, wherein the immune response comprises a B-cell response or a T cell response.

82. The method of claim 81, wherein the response is a T-cell response which comprises
30 generation of cytolytic T-cells specific for the host cells presenting the portion of the expression product of the nucleic acid molecule or cells of the subject expressing the human cancer associated antigen.

83. The method of claim 79, wherein the nucleic acid molecule is a NA Group 3 molecule.

84. The method of claims 79 or 80, further comprising treating the host cells to render
5 them non-proliferative.

85. A method for treating or diagnosing or monitoring a subject having a condition
characterized by expression of an abnormal amount of a protein encoded by a nucleic acid
molecule that is a NA Group 1 molecule, comprising

10 administering to the subject an antibody which specifically binds to the protein or a
peptide derived therefrom, the antibody being coupled to a therapeutically useful agent, in an
amount effective to treat the condition.

86. The method of claim 85, wherein the antibody is a monoclonal antibody.

15 87. The method of claim 86, wherein the monoclonal antibody is a chimeric antibody or a
humanized antibody.

88. A method for treating a condition characterized by expression in a subject of abnormal
20 amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid
molecule, comprising

administering to a subject a pharmaceutical composition of any one of claims 13-25
and 37-47 in an amount effective to prevent, delay the onset of, or inhibit the condition in the
subject.

25 89. The method of claim 88, wherein the condition is cancer.

90. The method of claim 88, further comprising first identifying that the subject expresses
in a tissue abnormal amounts of the protein.

30 91. The method of claim 89, further comprising first identifying that the subject expresses
in a tissue abnormal amounts of the protein.

92. A method for treating a subject having a condition characterized by expression of abnormal amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule, comprising

- 5 (i) identifying cells from the subject which express abnormal amounts of the protein;
- (ii) isolating a sample of the cells;
- (iii) cultivating the cells, and
- (iv) introducing the cells to the subject in an amount effective to provoke an immune response against the cells.

10

93. The method of claim 92, further comprising rendering the cells non-proliferative, prior to introducing them to the subject.

94. A method for treating a pathological cell condition characterized by aberrant expression of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule, comprising

15 administering to a subject in need thereof an effective amount of an agent which inhibits the expression or activity of the protein.

20 95. The method of claim 94, wherein the agent is an inhibiting antibody which selectively binds to the protein and wherein the antibody is a monoclonal antibody, a chimeric antibody or a humanized antibody.

96. The method of claim 94, wherein the agent is an antisense nucleic acid molecule

25 which selectively binds to the nucleic acid molecule which encodes the protein.

97. The method of claim 94, wherein the nucleic acid molecule is a NA Group 3 nucleic acid molecule.

30 98. A composition of matter useful in stimulating an immune response to a plurality of a proteins encoded by nucleic acid molecules that are NA Group 1 molecules, comprising a plurality of peptides derived from the amino acid sequences of the proteins, wherein

-269-

the peptides bind to one or more MHC molecules presented on the surface of the cells which express an abnormal amount of the protein.

99. The composition of matter of claim 98, wherein at least a portion of the plurality of
5 peptides bind to MHC molecules and elicit a cytolytic response thereto.

100. The composition of matter of claim 99, further comprising an adjuvant.

101. The composition of matter of claim 100, wherein said adjuvant is a saponin, GM-CSF,
10 or an interleukin.

102. The composition of matter of claim 98, further comprising at least one peptide useful
in stimulating an immune response to at least one protein which is not encoded by nucleic
acid molecules that are NA Group 1 molecules, wherein the at least one peptide binds to one
15 or more MHC molecules.

103. An isolated antibody which selectively binds to a complex of:
(i) a peptide derived from a protein encoded by a nucleic acid molecule that is a
NA Group 1 molecule and
20 (ii) and an MHC molecule to which binds the peptide to form the complex,
wherein the isolated antibody does not bind to (i) or (ii) alone.

104. The antibody of claim 103, wherein the antibody is a monoclonal antibody, a chimeric
antibody, a humanized antibody, or a fragment thereof.

SEQUENCE LISTING

<110> Ludwig Institute for Cancer Research

5 <120> BREAST, GASTRIC, AND PROSTATE
CANCER ASSOCIATED ANTIGENS AND USES THEREFOR

10 <130> L0461/7064WO

<150> US 60/136,526
<151> 1999-05-2815 <150> US 60/153,454
<151> 1999-09-10

<160> 1332

20 <170> FastSEQ for Windows Version 3.0

<210> 1
<211> 1231
<212> DNA
<213> Homo sapiens25 <400> 1

gagaggacca	agctaaacaa	gcatatgtt	acaagttgga	agaattaatg	aaaattggca	60
ctccagttaa	agttcggtt	caggaagctg	agaacggcc	aaaaatgtt	gaagaactag	120
gacagaggct	gcagcattat	gccaaagatag	cagctgactt	cagaataaag	gtggaaaat	180
acaaccatat	tgtatggatct	gaaatgaaaa	aagtggagaa	gtctgttaat	gaatgtatgg	240
aatggatgaa	taatgtcatg	aatgttcagg	ctaaaaaagag	tcttgatcag	gtcccgatgg	300
taatgtctca	ggaaaattaaa	acaaaaatca	aggaatttcaa	caacacatgt	gaaccgggtt	360
taacacaacc	gaaacccaaa	atggatcac	ccaaactgaa	aagaactcca	aatggcccaa	420
atattgataa	aaatggagaa	gatttagaa	acaaaaacaa	ttttgtgtct	gaacctccac	480
atcagaatgg	tgaatgttac	cctaattgaga	aaaattctgt	taatatggac	ttggactaga	540
taacccat	ttggcctatt	ccttcattt	ataaaaatatt	tttgcctat	tatgtgactc	600
tacataacat	actgaaaacta	tttatatttt	cttttttaag	gatattttaga	attttgtgt	660
attatgttgc	aaaagaaaaaa	aagctttaag	tctgtatct	ttatgtatct	aaaaggggaaa	720
attgccttgg	taactttcag	attctgtgg	aatttgtaat	tcataactaag	ctttctgtgc	780
agtctcncca	tttgcattcac	tgaggatgaa	actgacttt	gtcttttgg	aaaaaaaac	840
ttgtactgct	ttttcaagag	ggctgtgatt	aaaatctta	agcatttgg	cctgccaagg	900
tagtttctt	gcattttct	ctccatttcag	catgtgtgt	gtgtggatg	tttataaaca	960
agactaagtc	tgacttcata	agggtttct	aaaaccattt	ctgtccaaga	gaaaatgact	1020
ttttgttttg	atattaaaaa	ttcaatgagt	aaaacaaaaag	ctagtcaat	gtgttagcag	1080
45 catgcagaac	aaaaacttta	aactttctct	ctcnctatac	agtatattgt	catgtgaaag	1140
tgtggaatgg	aagaaatgtc	gatccgtt	taactgattt	tgaacacttt	tatgagctt	1200
aaaataaaagt	tcatctttag	gtgtcatttc	t			1231

50 <210> 2
<211> 965
<212> DNA
<213> Homo sapiens

55 ggagaatgaa	atgtcttctg	aagctgacat	ggagtgtctg	aatcagagac	caccagaaaa	60
cccagacact	gataaaaaatg	tccagcaaga	caacagtcaa	gctggAACAC	agccccaggt	120
acaaaactgt	gctcaacaaa	cctcacagtc	tcccccttca	cctgaactt	cctcagaaga	180
aaacaaaatc	ccagatgctg	acaaagcaaa	tgaaaaaaaa	gttgaccagc	ctccagaagc	240
taaaaagccc	aaaataaaagg	tggtaatgt	tgagctgcct	attgaagcca	acttggtctg	300
60 gcagtttaggg	aaagacccctc	ttaacatgt	tattgagaca	gagggtaa	tgataatgca	360
agataaaattt	aaaaagaaaa	ggaatgtatgc	aaaaatgca	gttgaggaa	atgtgtatga	420
gttcagagac	aagctgtgt	gaccatatga	aaaatttata	tgtgagcagg	atcatcaaaa	480
ttttttgaga	ctccctcacag	aaactgaa	ctggctgtat	gaagaaggag	aggaccaagc	540

<213> Homo sapiens

<400> 547

	gagacagagt	ctcaactctgt	cgctcaggct	ggagtgcagt	agtggatct	cggtcacct	60
5	gcctctaggg	ttcaagccat	tctctgcct	cagcctcca	agtagctggg	attacagagg	120
	cgtgcaccac	catgaccacc	taatttttc	ttgttggtt	tgttttttg	gtttttgttt	180
	tttttggana	tggagtctct	ctctgtcacc	caggctggag	tgcagtagtg	ggatctcagc	240
	tcacctgcct	ccagggttca	agccattctc	ctgcctcagc	ctcccaagta	gttgggattta	300
10	cagaggcgtg	caccaccatg	cccacctaatt	tttttcttgt	tttgntttc	tggttctttt	360
	tganacanag	tcttgctctg	tgcggcaagc	tggagtgcatt	tggcgcgatc	taagctca	420
	gcaagctcca	cctcccagat	tcacgcccatt	ctgctatgag	taaaatcacc	tggttggatt	480
	gaatgatgct	gctgcatcat	canggtagaa	tgggaatctt	gtgcagactt	tatgcattc	540
	gtgaacaaaq	taattttca	cactggatt	ttatgtctgt	aaggcttca		589

15 <210> 548

<211> 776

<212> DNA

<213> Homo sapiens

<400> 548

agatactccc atcagaatcc aaacaaaagg actatgaaga aagtcttgg gattctgaga
gtctctgtga gactgttca cagaaggatg tgcgttacc caaggctgcg catcaaaaag
aaatagataa aataaatggc aaattagaag ggtctctgt taaagatggc cttctgaagg
ctaactgcgg aatgaaagtt tctattccaa ctaaagcctt agaattgtat gacatgaaa
ctttcaaagc agagectccc gagaagccat ctgccttcga gcctgcatt gaaatgaaa
agtctgtcc aaataaagcc ttggaattga agaatgaaca aacattgaga gcagatgaga
tactcccatc agaatccaa caaaaggact atgaagaaag ttcttggtat tctgagatc
tctgtgagac tgttcacag aaggatgtt gtttacccaa ggctgcgtat caaaaagaaa
tagataaaaat aaatggaaaa tttagaagggt ctccctttaa agatggctt ctgaaggcta
actgcggaat gaaagttct attccaacta aagccttaga attgatggac atgc当地
tcaaagcaga gcctccccag aagccatctg cttcgagcc tgccattgaa atgcaaaaag
tctgttncaa ataaaagcctt ggnattggaa gaatgaacaa acattgagag ccgatgaaa
actnccatca gaatnccaac caaanggctn ttgaagaaaag ttcttggtat ttttga

35 <210> 54

<211> 820

<212> DNA

<213> Homo sapiens

<400> 549

gaatgaacaa	acattggagag	cagatgagat	actcccatca	gaatccaaac	aaaaggacta	60
tgaagaaagt	tcttgggatt	ctggagact	ctgtgagact	gtttcacaga	aggatgttg	120
tttacccaag	gctacacatc	aaaaagaaat	agataaaata	aatggaaaat	tagaagagtc	180
tcctgtataat	gatggtttc	tgaaggctcc	ctgcagaatg	aaagtttcta	ttccaaactaa	240
agccttagaa	ttgatggaca	tgcaaacttt	caaagcagag	cctcccgaga	agccatctgc	300
cttcgagcct	gccattgaaa	tgcaaaagtc	tggttccaaat	aaagccttgg	aattgaagaa	360
tgaacaaaca	ttgagagcag	atcagattt	cccttcagaa	tcaaaaacaaa	agaaggttga	420
agaaaattct	tgggattctg	agagtctccg	tgagactgtt	tcacagaagg	atgtgtgtgt	480
cccaaggctca	cacatcaaaa	agaaatggat	aaaataagtg	gaaaattaga	gagactctta	540
gtttcatcct	gtaccaagag	agtctgatag	gtccccctcat	tttttcttcc	tgtgtcacgt	600
tttggaatct	tgtcaagcan	gaagctgaac	atttottaaaa	tcactaccta	tcaaaatctt	660
gatcngtcan	tcttgggaaag	accangggac	ttcaaaaagat	actgggaccc	ctcnccggaaa	720
tggccnatg	aaaaaaaaagt	tttgttcttga	aaaaactntc	gaancaaana	attaatcccc	780
ttttqgnacc	aaaaagttaat	ngggccaaac	ccttccgggg			820

55

<210> 550

<211> 713

<212> DNA

	tacagaggac atagcagtat taaggataa tgaagtcaca gcttcagacg ctccatcctt	180
	tcttagcaa gttagctcta cttgtatctg ttctgttta tataatatgg ttgcataa	240
	ctgttttaa aaaaagttct gttttcaaa aaaattttaa gctatgaaaa tcactgatta	300
5	agtcacacc tcatttaca aaagaggca cacaactca gagcacttat gcctcaccat	360
	aggtcacaaa gccaagtagc tccaggccag aaatggctt taggtctcc gtctganact	420
	ggcatttgat gccagtatt tcctccatat ttaggagaaa tgtatagatt taaaatata	480
	actcanagaa aatgcatgca tatggaaaa ctgtttcta gctcaatcat gtttagggc	540
	aacaggtaa aaaaatagnic tccagctaaa ttgngaatnc taaaattaga aatacaaagt	600
10	aggccataga aactaatatc ntaacatcg gggagangga ggcaattaat ctattgtaa	660
	agaacaagtt tggtaaaaaac tccanaaaaa aggacctcg aagggagaaa aat	713
	 <210> 551	
	<211> 708	
	<212> DNA	
15	<213> Homo sapiens	
	 <400> 551	
	caaaaaggact atgaagaaaag ttcttggat tctgagagtc tctgtgagac tgtttcacag	60
	aaggatgtgt gtttacccaa ggctgcgcatt caaaaagaaa tagataaaaat aaatggaaaa	120
20	tttagaagggt ctccctgttaa agatggctt ctgaaggcta actgcggaaat gaaagttct	180
	atcccaacta aagccttaga attgatggac atgcaaaactt tcaaagcaga gcctcccgag	240
	aagccatctg ccttcgagcc tgccattgaa atgcaaaagt ctgtttccaaa taaagccttg	300
	gaattgaaga atgaacaaac attgagagca gatgagatc tcccatcaga atccaaacaa	360
	aaggactatg aagaaaggcc ttggattct gagactctt gtgagactgt ttacacagaag	420
25	gatgtgtgtt tacccaaaggc tgccatcaa aaagaaatag ataaaataaa tgaaaattta	480
	gaagggtctc ctgttaaaga tggcttctga aggctaactg cggaatgaaa gtctattca	540
	actaaaggct tagaattgtt ggacatgcca actttcaagc agacctccga gaagcctctg	600
	cttcnacctg cattgaatgc caaaatctgt tccaataaaag ctttggattt gaagaatgaa	660
	ccaacatgg agatctccct cagaatccac caaaggct	708
30	 <210> 552	
	<211> 716	
	<212> DNA	
	<213> Homo sapiens	
35	 <400> 552	
	catattgata tatacggttt tttaaatggg tattgttaatt aaatatctcc tcatttttct	60
	cttttaggag atgatgttc attttctct caagaaaatg aatatacaatt gtatcttgc	120
	tttgggtgtc agctttctta tgcgtatgaa ctaattgtcg ttgaagccac atattttgc	180
40	tttgttagttg aaataatcc tgcatttttttactccctgc ttcaatgtgt ttgttcacat	240
	tatcttgc gttttgatac atgtgttcag cttcccttcat ttgacactgt gtttacgtt	300
	ggtccttttg tgcatttttgc gaaaccaatg tttttctct tanagcatct ccggcataat	360
	tgagattaaat ttttaggctt ttggattttcc tttaggtttc anaaagtggt tgatggagca	420
	cctcatttttgc atatatcgta ctactccat caacattcat ttttcttgc aaacaagcat	480
45	ctcctgcaat gtggaaagca ggttcttgac tttttcttgc tgcacaatt tgatcatggg	540
	cttgcctcaggc agaagccaggc ttagatgggn ggtgattcaa tttctgcctc taatatttct	600
	ttggcttgggt tttccttcaa tttaaaaatg gaggatttng gttctcaact ntcanaaactt	660
	taagctgncc acttttttgc gaagccccctt ttagtaaaga atccttttgc agtttta	716
50	 <210> 553	
	<211> 713	
	<212> DNA	
	<213> Homo sapiens	
55	 <400> 553	
	aatgaacaaa cattgagagc agatggagata ctcccatcag aatccaaacaa aaaggactat	60
	gaagaaaagtt ctgggattc tgagactctc tgcgtatgactg tttcacagaa ggatgtgtgt	120
	ttacccaaagg ctgcgcattca aaaagaaaata gataaaaataa atggaaaattt agaagggtct	180
60	cctgtttaaag atggcttctt gaaggctaac tgcggatgaa aagtttctat tccaactaaa	240
	gccttagaat tgatggacat gcaaaacttcc aaagcagagc ctcccggaa gccatctgcc	300
	ttcgagctcg ccattgaaat gcaaaaatgtct gttccaaata aagccttggg attgaagaat	360
	gaacaaaacat tgagagcaga tgagataactc ccatcagaat ccaaacaaaa ggactatgaa	420
	gaaagttctt gggattctga gactctgtt gactgttt cacanaaggg tgcgtgttcc	480

<210> 557
 <211> DNA
 <212> DNA
 <213> Homo sapiens

5 <400> 557

gtttttttt	ttttgatacg	gagtctcgct	ctgtcgccca	ggctggattg	60
aatgatgtct	gctgcacat	caggtagat	ggaatcctg	tgcagctta	120
gaacaagtaa	ttcttcacac	tgttattttt	tgtctgtagg	cttcaggaan	180
tttttcttct	aaaataaga	ttctgttgac	acanactatt	ggtaagattt	240
gngatgctag	gactggcctc	ctagcatgag	ttgtgagtaa	agatctggc	300
ccaaaanaag	tttcttaactg	cttgctctc	atgagtttc	tgtttctgt	360
catattgata	tatacgggtt	ttaaatggt	tattgttaatt	aaatatctcc	420
cttttaggag	atgatgttg	atttcctct	caagaaaaatg	aatatcaatt	480
tttttgttgc	agcttctta	tgtgcacat	ctaattgctg	ttgaagccac	540
ttttagttg	aaataatttc	tgatctanag	actcctgctg	ttcaatgtgt	600
tatcttgttc	gtttgaaac	atgtttcan	cttccttcat	ttgacactgg	660
ggcttttgt	gcatgttctg	aaaccatgt	ntttcttctt	aaaaccatct	720
tgaaaataat	ttttaggctt	tttgggattt	ncttttgagc	ttcaaaaaag	770

20 <210> 558
 <211> 735
 <212> DNA
 <213> Homo sapiens

25 <400> 558

gaaagagtt	gaagtgaaac	aacaacttga	acaggctctc	agaatacaag	60
gaagagtgt	aaaagtatt	tgaatcaggt	ttctcacact	catgaaaatg	120
cttacatgaa	aattgcatt	tgaaaaagga	aattgccatg	ctaaaactgg	180
actgaaaacac	caataccagg	aaaaggaaaa	taaataacttt	gaggacatta	240
agaaaagaat	gctgaacttc	agatgaccct	aaaactgaaa	gaggaatcat	300
ggcatctcaa	tatagtggc	agcttaaagt	tctgatagtc	gagaacacaa	360
taaattgaag	aaaaaacaag	acaaagaaat	actagaggca	gaaattgaat	420
tagactggct	tctgctgtac	aagaccatga	tcaaatttg	cacaccatcc	480
acctgcttcc	cacattgcag	gagatgctt	tttgc当地	taactaaaag	540
tagtaccat	atataacaat	gaggngctcc	atcaaccact	tgctgaagct	600
ncaaaaggct	aaaaattaaat	ctcaattatg	cccagatgc	caaangaaaat	660
tttcagaaca	tgcccaagag	gaccacgtgg	aacaaagtgt	atacattgg	720
ccccatgttcc	aaacc				735

40 <210> 559
 <211> 664
 <212> DNA
 <213> Homo sapiens

45 <400> 559

ctgccttcga	gcctgcatt	gatatgcaaa	agtctgttcc	aaataaagcc	60
agaatgaaca	aacattgaga	gcagatgaga	tactccatc	agaatccaaa	120
atgaagaaag	ttcttggat	tctgagatc	tctgtgagac	tgtttcacag	180
gtttacccaa	ggctgcgcatt	caaaaagaaa	tagataaaaat	aatggaaaa	240
ctccctgttaa	agatggctt	ctgaaggcta	actgcggaaat	gaaagttct	300
aaggccttaga	attgatggac	atgcaaaactt	tcaaagcaga	gcctcccgag	360
ccttcgagcc	tgccattgaa	atgcaaaagt	ctgttccaaa	taaagccttg	420
atgaacaaac	attgagagca	gatgagatac	tcccattcaga	atccaaacaa	480
aagaaagtcc	ttgggattct	gagatctct	gtgagactgt	ttccagaagg	540
acccaangct	tgcgcntcaa	aaagaaatag	ataaaataaa	tggaaaatta	600
atgctgtga	atttaaaaca	ttcttgc当地	tgataaggc	ttctgttaaa	660
tgaa					664

60 <210> 560
 <211> 636
 <212> DNA
 <213> Homo sapiens

